

On the Range of Equilibria Utilities of a Repeated Epidemic Dissemination Game with a Mediator

Xavier Vilaça
INESC-ID, Instituto Superior Técnico
Universidade de Lisboa, Portugal
xvilaca@gsd.inesc-id.pt

Luís Rodrigues
INESC-ID, Instituto Superior Técnico
Universidade de Lisboa, Portugal
ler@ist.utl.pt

ABSTRACT

We consider eager-push epidemic dissemination in a complete graph. Time is divided into synchronous stages. In each stage, a source disseminates ν events. Each event is sent by the source, and forwarded by each node upon its first reception, to f nodes selected uniformly at random, where f is the fanout. We use Game Theory to study the range of f for which equilibria strategies exist, assuming that players are either rational or obedient to the protocol, and that they do not collude. We model interactions as an infinitely repeated game. We devise a monitoring mechanism that extends the repeated game with communication rounds used for exchanging monitoring information, and define strategies for this extended game. We assume the existence of a trusted mediator, that players are computationally bounded such that they cannot break the cryptographic primitives used in our mechanism, and that symmetric ciphering is cheap. Under these assumptions, we show that, if the size of the stream is sufficiently large and players attribute enough value to future utilities, then the defined strategies are Sequential Equilibria of the extended game for any value of f . Moreover, the utility provided to each player is arbitrarily close to that provided in the original game. This shows that we can persuade rational nodes to follow a dissemination protocol that uses any fanout, while arbitrarily minimising the relative overhead of monitoring.

Categories and Subject Descriptors

C.2.4 [Computer-Communication Networks]: Distributed Systems; K.6.0 [Management of Computing and Information Systems]: General—*Economics*

General Terms

Theory, Economics, Reliability

Keywords

epidemic dissemination, repeated games, monitoring

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICDCN '15 Jan. 4-7, Goa, India

Copyright 2015 ACM 978-1-4503-2928-6 ...\$15.00.

1. INTRODUCTION

This paper addresses the impact of rational behaviour in epidemic dissemination protocols, executed over a complete graph [6]. An epidemic dissemination protocol operates as follows: a source splits a stream of bits into ν events, which are sent to a set of f nodes chosen uniformly at random, where f is known as the *fanout*; nodes repeated this procedure upon the first reception of each event. Protocols of this type are known to achieve a good tradeoff between high reliability of event delivery and communication overhead, and have been used in a variety of applications such as video streaming [15, 14]. In this context, rational behaviour may be characterised by the aim of maximising a utility, given as follows. Rational nodes value the stream, but they prefer to send as few messages as possible, in order to spare bandwidth. Therefore, the utility can be defined as the difference between the benefits, which increase with the number of received events, and the communication costs for sending messages. This setting poses the problem that rational nodes always prefer not to forward any messages.

To address this issue, we explore the possibility of nodes interacting repeatedly in multiple executions of the dissemination protocol, in order to hold nodes accountable for their present behaviour by adjusting their utility in the future. This models periodic streaming sessions (e.g., weekly sporting events). Using Game Theory [17], we study incentives to persuade rational nodes to follow the protocol, also admitting the possibility that some nodes may be acquiescent, i.e., obedient to the protocol [3, 20]. We assume that rational nodes do not collude. We model interactions as an infinitely repeated game where future utilities are discounted to the present by some factor $\delta \in (0, 1)$, which determines the value given by players to future utilities¹. This is adequate when players are uncertain about the number of future interactions and value utilities obtained in the present over future ones [17]. Our aim is to study the range of the values of f used by dissemination protocols that correspond to equilibria strategies of the repeated game, that is, where no player has any incentive to deviate from the protocol assuming that other players also do not deviate.

When players interact repeatedly, incentives may be based on *direct* or *indirect* reciprocity [18]. With *direct reciprocity*, each player i adapts his strategy towards every other player j in reaction to past actions of j directed towards i only. An

¹We use the designations player and node to describe the entities of the system: node refers to the operational side of the entity; player refers to the (rational) user controlling the node.

example is the tit-for-tat strategy [4, 7]. Most epidemic dissemination systems that cope with rational behaviour also use direct reciprocity [15, 14, 12]. Unfortunately, strategies of this type are vulnerable to the redundancy of epidemic dissemination [19]. Namely, if a player i does not cooperate with j , then at best j may punish i . However, when redundancy is high, i has other neighbours from whom i receives events. Hence, the impact of such punishment is arbitrarily low, since i continues to receive events with sufficiently high probability. This makes direct reciprocity an ineffective type of incentives when f is large.

Indirect reciprocity circumvents the limitations of direct reciprocity by having nodes sharing information regarding private observations. This allows all the nodes to coordinate on effective punishments against any player i by not forwarding any event to i , decreasing his utility to 0. In Game Theoretical terms, we can punish each player by decreasing his utility to the minimax value. Under this possibility, we might apply a well known set of results called Folk Theorems, which state that we may devise equilibria strategies for infinitely repeated games that provide any feasible strictly positive utility to every player [16], given that δ is sufficiently large. In particular, these results imply that it is possible to sustain cooperation while using any fanout for disseminating events. However, existing proofs of Folk Theorems assume that some underlying monitoring infrastructure provides information about the behaviour of each node, at no cost to the players [16]. Such assumption is unrealistic, since any implementation of a monitoring mechanism always incurs communication costs.

Goal : We show that cooperation can be sustained for any fanout using a monitoring mechanism that is not free of cost. For this purpose, we prove the existence of equilibria strategies for the game induced by the monitoring mechanism. We consider the notion of Sequential Equilibrium [13], which is stronger than Nash Equilibrium (NE) since it excludes strategies that rely on non-credible threats. For instance, when considering monitoring, the notion of NE does not evaluate the optimality of a strategy when a player has observed a deviation and has to communicate this fact to other players. In line with the Folk Theorems, we also aim at providing any feasible and strictly positive utility to each player, which requires the minimisation of the communication overhead of monitoring relative to the original dissemination protocol. Existing works have faced the challenge of implementing a distributed monitoring mechanism [9], and performed a game theoretical analysis of epidemic dissemination [19]. To the best of our knowledge, none has studied the range of f used by equilibria strategies of the repeated epidemic dissemination game.

Challenges: To fulfil our goal, we define a monitoring mechanism that extends the infinitely repeated epidemic dissemination game with additional communication rounds, and we propose a set of strategies for this game. In addition to disseminating a stream of events, nodes review the behaviour of every player and share this information, which is used to decide when to punish each player. When a punishment against a player i is in place, no node sends events to i , denying any benefit to i . This way, the threat of punishment out-weighs the gain from deviating from the specified strategy. The following main challenges are addressed:

Challenge 1: Strategic Monitoring. Players may deviate

when sharing monitoring information. For instance, a punishment of any player i causes the overall reliability of dissemination to decrease. Hence, players are not willing to share information incriminating i .

Challenge 2: Mixed strategies. Nodes randomly select the neighbours to forward each event. The difficulty lies in preventing players from biasing this selection.

Challenge 3: Hidden Events. While observing the actions of i , player j does not observe the set of events received by i . Thus, the monitoring mechanism may raise false positives, causing players to be punished undeservedly.

Challenge 4: Overhead of monitoring. Nodes cannot share information regarding each disseminated event, otherwise the overhead of monitoring is not minimised. Though, monitoring only a subset of disseminated events introduces the problem of false negatives, where misbehaviour is undetected.

Summary of Contributions: We devise a monitoring mechanism and a set of strategies that are Sequential Equilibria of the extended game for any fanout, provided that ν and δ are sufficiently large. These strategies also minimise the communication overhead of monitoring relative to the original dissemination protocol. It is important to notice that the overhead is minimised only relative to the total size of the stream. We use symmetric cryptography to cope with strategic monitoring. Assuming that symmetric ciphering costs are negligible and players are computationally bounded, this allows cheap punishments to be applied to players, while creating incentives for them to continue forwarding events. Our results offer an improvement over existing work towards the goal of designing a practical monitoring mechanism [9], which does not consider strategic monitoring. We use a pseudo-random number to address the challenge raised by mixed strategies. We address the challenge of hidden events by relying on a trusted mediator to collect information about the events sent to and received from each node i . We minimise monitoring overhead by having nodes reporting in expectation on only a sub-linear number of events. The mediator coordinates the selection of these events to ensure that the probability of false negatives is sufficiently low, and that no false positives are raised.

Paper Organisation: The remainder of the paper is organised as follows. In Section 2, we introduce our model. Section 3 contains the monitoring mechanism and strategies. In Section 4, we perform a Game Theoretical analysis of the strategies to prove the main result. Section 5 concludes the paper with a short discussion.

2. MODEL

We consider a synchronous message passing system with reliable and authenticated communication. The set of nodes is denoted by \mathcal{N} , and n is its cardinality. We consider that \mathcal{N} is common knowledge². Players do not collude and have perfect recall. Time is divided into stages, which are further divided into τ synchronous rounds. In each stage t , the source disseminates a set \mathcal{E}^t of ν events, drawn from a much larger but finite set \mathcal{E} . The process of generating \mathcal{E}^t must be sufficiently random, such that any player i can guess any $e \in \mathcal{E}^t$ beforehand only with a sufficiently small probability. This is to justify the assumption that players prefer to receive

²Every player knows this set, knows that every player knows this set, and so on.

these events than to try to guess them. Every event $e \in \mathcal{E}^t$ has a unique identifier $id \in \{1 \dots \nu\}$, and is disseminated as follows: first, the source sends e to a subset of f neighbours chosen uniformly at random, where f is the fanout. Then, each node forwards e upon its first reception, also to a subset of f neighbours selected uniformly at random. This process ends until no node forwards e or until a maximum delay of ρ rounds to deliver e is reached, after which the event is said to expire. We assume that ρ does not increase with ν . A different event is introduced in each round by the source. We let $\tau \geq \nu + \rho$ to allow every event to be disseminated until it expires. We consider that the source is obedient to the protocol and may act as a trusted mediator. Using the language of [3, 20], we say that the source is acquiescent. We treat every other player as rational, although our results still hold if other acquiescent players exist. In Section 5, we discuss how to distribute the role of the mediator. Players are assumed to be computationally bounded - they cannot break the cryptographic primitives used in our strategies, in the time required by a stage.

Interactions are modelled as an infinitely repeated epidemic dissemination game. An action of player i is a vector $\vec{a}_i \in \mathcal{A}_i$ specifying for each $j \in \mathcal{N} \setminus \{i\}$ a message $a_i(j)$ sent by i to j . Messages contains a finite number of tuples (id, e) , where $id \in \{1 \dots \nu\}$ and $e \in \mathcal{E}$. A history $h \in \mathcal{H}$ is a finite sequence of action profiles $\vec{a} \in \mathcal{A}$ specifying all the messages sent after multiple rounds. Any player i cannot completely observe a history; instead, i only observes a corresponding private history $h_i \in \mathcal{H}_i$, specifying the messages sent and received by i in h . We say that histories h_i and h are from stage t if the first round following these histories belongs to t . A strategy $\sigma_i \in \Sigma_i$ specifies a probability distribution $\sigma_i(\cdot|h_i)$ over the actions taken by i in the round immediately succeeding the observation of h_i . A strategy profile $\vec{\sigma} \in \Sigma$ specifies the strategy followed by every player. Given any $\vec{\sigma} \in \Sigma$, each player i forms a belief μ regarding the realised history $h \in \mathcal{H}$ after the observation of $h_i \in \mathcal{H}_i$, in the form of a probability $\mu^{\vec{\sigma}}(h|h_i)$. μ is also common knowledge. If we fix some strategy profile $\vec{\sigma}$ that is expected to be followed by any player, then in any realised history $h \in \mathcal{H}$ the behaviour of an acquiescent player i is always compatible with i having followed $\vec{\sigma}$ in h . Henceforth, whenever referring to a history $h \in \mathcal{H}$ and strategy profile $\vec{\sigma} \in \Sigma$, we will implicitly consider that the behaviour of the source (and the mediator) in h is compatible with $\vec{\sigma}$. Thus, we never analyse histories where acquiescent entities have deviated in the past, since this analysis is irrelevant to the proof of equilibrium. This fact is also captured by the defined belief system: for any $h_i \in \mathcal{H}_i$ and $h \in \mathcal{H}$, we have $\mu^{\vec{\sigma}}(h|h_i) > 0$ only if all acquiescent entities have followed $\vec{\sigma}$ in h . This implies that i never observes a deviation of some acquiescent entity in any h_i .

2.1 Utility

The expected utility of the infinitely repeated game obtained by any player i depends on the following factors: i) the private history $h_i \in \mathcal{H}_i$ observed by i , initially equal to \emptyset ; ii) the strategy profile $\vec{\sigma}$ followed by every player; iii) the belief system μ ; and iv) the realised utility u_i of receiving events and sending messages during each stage following the observation of h_i . The pair $(\vec{\sigma}, \mu)$ allows i to form an expectation of what occurred in the past, in terms of a probability distribution over the histories $h \in \mathcal{H}$. Given any h , i can predict the future behaviour of any player j in any future

stage t , given that j follows σ_j . More precisely, $\vec{\sigma}$ defines a probability distribution over the outcomes $\mathcal{Z}^t \subseteq \mathcal{H}$ of stage t , where each outcome $z \in \mathcal{Z}^t$ is a history that matches the end of stage t . We denote by $P^{\vec{\sigma}}(z|h)$ the probability of z being reached after h , given that players follow $\vec{\sigma}$. The realised utility for stage t is a function $u_i(z)$ of the outcome $z \in \mathcal{Z}^t$ reached in stage t . This function quantifies the average benefits of receiving events and the average costs of sending messages, per disseminated event. More precisely, we consider that every player receives a benefit β per received event $e \in \mathcal{E}^t$ and incurs a cost α per bit sent in a message. $u_i(z)$ is given by the total benefits minus the total costs. We divide $u_i(z)$ by ν in order to normalise it to the average utility per disseminated event. We provide a formal definition of $u_i(z)$ in Section 3.4, after defining the monitoring mechanism.

With this in mind, the expected utility for stage t given the observation of h_i , denoted by $\mathbb{E}^{\vec{\sigma}, \mu, t}(u_i|h_i)$, is the weighted sum over every history h compatible with the observation of h_i and outcomes of stage t following h :

$$\mathbb{E}^{\vec{\sigma}, \mu, t}(u_i|h_i) = \sum_{h \in \mathcal{H}} \sum_{z \in \mathcal{Z}^t} \mu^{\vec{\sigma}}(h|h_i) P^{\vec{\sigma}}(z|h) u_i(z). \quad (1)$$

Finally, for any $i \in \mathcal{N}$ and $h_i \in \mathcal{H}_i$ from any stage t , $\mathbb{E}^{\vec{\sigma}, \mu}(u_i|h_i)$ is the weighted infinite sum over every stage $t' \geq t$ of the expected utility of t' . We use a discount factor $\delta \in (0, 1)$ to discount future utilities to the present:

$$\mathbb{E}^{\vec{\sigma}, \mu}(u_i|h_i) = \sum_{t' \geq t} \delta^{t'-t} \mathbb{E}^{\vec{\sigma}, \mu, t'}(u_i|h_i). \quad (2)$$

2.2 Notion of Equilibrium

We consider the notion of Sequential Equilibrium (SE) [13]. We say that a pair $(\vec{\sigma}, \mu)$ is a SE if it is Sequentially Rational and Consistent. $(\vec{\sigma}, \mu)$ is Sequentially Rational if σ_i maximises $\mathbb{E}^{\vec{\sigma}, \mu}(u_i|h_i)$ for any $h_i \in \mathcal{H}_i$, conditional on the belief that other players follow $\vec{\sigma}$ after h_i is observed. Formally, let $(\sigma'_i, \vec{\sigma}_{-i})$ be the strategy profile where every $j \in \mathcal{N} \setminus \{i\}$ follows $\sigma_j \in \vec{\sigma}_{-i}$ and i follows σ'_i .

DEFINITION 2.1. *The pair $(\vec{\sigma}, \mu)$ is Sequentially Rational iff for every $i \in \mathcal{N}$, $h_i \in \mathcal{H}_i$, and $\sigma'_i \in \Sigma_i$, it holds:*

$$\mathbb{E}^{\vec{\sigma}, \mu}(u_i|h_i) \geq \mathbb{E}^{(\sigma'_i, \vec{\sigma}), \mu}(u_i|h_i).$$

We do not include the formal definition of Consistency (c.f. [13]). Informally, $(\vec{\sigma}, \mu)$ is Consistent if μ is defined using the Bayes rule according to the behaviour specified by $\vec{\sigma}$, whenever possible. When some h_i is observed that is inconsistent with the hypothesis that players have been following $\vec{\sigma}$, the definition of consistency requires the specification of an alternative hypothesis for explaining the observed behaviour. We fix a belief system μ that suits our purposes, defined as follows. For any $\epsilon > 0$, define $\vec{\sigma}^\epsilon$ as the strategy profile where every i follows $\sigma_i(\cdot|h_i)$ after each $h_i \in \mathcal{H}_i$ with probability $1 - \epsilon$, and, with probability ϵ , i follows any available action with positive probability. Since every history is consistent with players following $\vec{\sigma}^\epsilon$, we can apply the Bayes rule to completely define $\mu^{\vec{\sigma}^\epsilon}$. Then, we set $\mu^{\vec{\sigma}}(\cdot|h_i) = \lim_{\epsilon \rightarrow 0} \mu^{\vec{\sigma}^\epsilon}(\cdot|h_i)$. The intuition is that the observed behaviour is explained by players following $\vec{\sigma}$ and occasionally making mistakes with a small probability.

The above definition of Sequentially Rational pair $(\vec{\sigma}, \mu)$ is problematic, since it requires the analysis of all possible

alternative strategies to that specified by $\vec{\sigma}$. Fortunately, we can simplify this task by analysing only local deviations according to the One-deviation Property [10]. A local deviation for player i after private history $h_i \in \mathcal{H}_i$ is an action that is not prescribed by σ_i with positive probability, while every player still follows $\vec{\sigma}$ after i observes h_i , and i follows σ_i in every round following the deviation. Formally, let $\mathbb{E}^{\vec{\sigma}, \mu}(u_i | h_i, \vec{a}_i)$ denote the expected utility of i when every player follows $\vec{\sigma}$ after h_i is observed, except only that i follows \vec{a}_i immediately after observing h_i .

PROPOSITION 2.2. One-deviation [10]. *The pair $(\vec{\sigma}, \mu)$ is Sequentially Rational iff for every $i \in \mathcal{N}$, $h_i \in \mathcal{H}_i$, and $\vec{a}_i, \vec{a}_i' \in \mathcal{A}_i$ such that $\sigma_i(\vec{a}_i | h_i) > 0$, it holds:*

$$\mathbb{E}^{\vec{\sigma}, \mu}(u_i | h_i, \vec{a}_i) \geq \mathbb{E}^{\vec{\sigma}, \mu}(u_i | h_i, \vec{a}_i').$$

2.3 Central Claim

In line with Folk Theorems [16], we aim at defining equilibria strategies for any fanout f , while providing to each player any strictly positive expected utility of the dissemination game as the average utility of the repeated game. The average utility is computed as $(1 - \delta)\mathbb{E}^{\vec{\sigma}, \mu}(u_i | \emptyset)$, where \emptyset is the initial empty history. When all players forward events using a fanout f , the expected utility for any stage is $\bar{u}(f) = q(f)(\beta - \gamma f)$, where $q(f)$ is the probability of a given node receiving each event, β is the benefit per event, and γ is the cost for forwarding the event. Folk Theorems imply that, if $\beta > \gamma f$ and δ is sufficiently large, then an equilibrium strategy for the infinitely repeated game exists that yields $\bar{u}(f)$ to every player as the average utility of the repeated game. In this work, we aim to prove a slightly weaker result. First, we can only ensure the existence of such strategy if $\beta > c\gamma f$ for some constant $c > 0$ that may be greater than 1³. Second, we need both δ and ν to be sufficiently large. Finally, due to the overhead of monitoring, we can only provide to each player a utility arbitrarily close but never exactly equal to $\bar{u}(f)$. Theorem 2.3 formalises the central claim.

THEOREM 2.3. *Fix any fanout f and constant $\epsilon > 0$. There exist a monitoring mechanism, a strategy profile $\vec{\sigma}$ for the dissemination game, and a belief system μ , and there exist constants $c > 0$, $\bar{\delta} \in (0, 1)$, and $\bar{\nu} \in \mathbb{N}$, such that, if $\beta > c\gamma f$, then, for every $\delta \in (\bar{\delta}, 1)$ and $\nu > \bar{\nu}$, $(\vec{\sigma}, \mu)$ is a SE and, for every $i \in \mathcal{N}$:*

$$|(1 - \delta)\mathbb{E}^{\vec{\sigma}, \mu}(u_i | \emptyset) - \bar{u}(f)| < \epsilon.$$

The core of the proof of this theorem is the definition of a monitoring mechanism executed every stage that, combined with a dissemination strategy, provides a SE. Also, its costs are sub-linear on ν . By dividing by ν , we obtain the average monitoring costs per event. For an arbitrarily large ν , the impact of monitoring in the expected utility is arbitrarily small. Alternatively, we could fix ν and minimise the overhead of monitoring by only executing the mechanism once every period of t stages. Here, we could arbitrarily increase t to decrease the impact of monitoring. This alternative approach would be almost identical to the one used in this paper, both in terms of the definition of the monitoring mechanism and the main arguments of the proof, albeit its greater complexity. We discuss this possibility in Section 5.

³With appropriate optimisations, we can get $c \leq 3$.

3. MONITORING MECHANISM

We extend the dissemination game by adding τ^M monitoring rounds to each stage. For convenience, these rounds are added at the beginning of the stage. Therefore, a stage is now divided into $\tau = \tau^M + \tau^D$ rounds: the first τ^M rounds are used for exchanging monitoring information; the last τ^D rounds are used to disseminate events as in the dissemination game. Recall that ρ is the maximum delay to deliver any event. We still set $\tau^D \geq \nu + \rho$ to allow every event to be disseminated until it expires. The trusted mediator is responsible for collecting monitoring information from each node. This consists in *accusations* and *reports*. An accusation flags a deviation of some player, whereas a report, which is relative to some $id \in \{1 \dots \nu\}$, indicates for each pair of nodes (i, j) whether i sent to/received from j a tuple containing id . For each $f > 0$, we define a strategy profile $\vec{\sigma}^f$, and fix a belief μ^f that is Consistent with $\vec{\sigma}^f$ as defined in [13]. The main incentive for players to not deviate is based on indirect reciprocity. The mediator collects accusations and reports in stage t relative to the behaviour of each player i in stage $t - 1$, and gives a verdict on whether i should be punished during stage t . If i is punished, then i is still obliged to incur costs in stage t ; if i deviates again, then his punishment is extended to stage $t + 1$. We address the main challenges as follows:

Strategic Monitoring: We enforce the following two properties. First, each player i is only allowed to send monitoring information relative to other nodes; information sent by i has no effect on the probability of i being punished. Second, regardless of the punishments being applied, the reliability of dissemination remains constant. We achieve this using commutative symmetric ciphering. For each node i , the mediator sends a key κ_i to every node $j \neq i$, used by j to cipher events sent to i . While being punished, i is unable to retrieve the disseminated events. However, i still forwards each ciphered event normally. Every node $j \neq i$ receiving such event is aware that i is being punished and is capable of retrieving the original event, given his knowledge of κ_i .

Mixed Strategies: In dissemination rounds, every node i forwards each received event to a set S of f nodes. To prevent players from biasing the selection of S , S is specified by a pseudo-random number generator (PRNG), seeded by a random number sd_i , sent by the mediator to i only.

Hidden Events: For each pair of nodes (i, j) and $id \in \{1 \dots \nu\}$, i reports to the mediator the round when i first received from/sent to j a tuple containing id . Given these reports, the mediator is able to determine whether i received id and forwarded it as expected. More precisely, the mediator verifies whether i forwarded id as specified by PRNG and sd_i , immediately after its first reception. For each identifier id not received by i , the mediator verifies if i did not send a tuple with identifier id . If any of the above conditions fails, then the mediator triggers a punishment of i .

Monitoring Overhead: Nodes report only on a subset of events selected by the mediator in a non-deterministic fashion. The expected number of reported events is sub-linear on ν , such that the relative overhead of monitoring decreases as ν increases. More precisely, we split the identifier space $\{1 \dots \nu\}$ into n^S sequences, each containing n^E different identifiers. We define $\tau^M = 2n^S + 2$. In the first monitoring round, nodes send accusations to the mediator. In each even round $r < \tau^M$, for each node i , the mediator decides whether to monitor the identifiers from the sequence corresponding

to r , with independent probability p^* . If so, then the mediator sends a notification to every node $j \neq i$, in which case j must report, in round $r + 1$, on the identifiers from the corresponding sequence. In the last round, the mediator sends the seeds, the keys, and the verdicts on each player. These parameters are specified in Section 4.3.

We now describe the main components of the monitoring mechanism in more detail.

3.1 Symmetric Cipher

Let \mathcal{K} be the set of keys. For any key $\kappa \in \mathcal{K}$, let $(v)_\kappa$ be the operation that returns datum v ciphered with κ . We will consider primitives for which the ciphering operation is the same as that of deciphering. Hence, we consider that it holds $((v)_\kappa)_\kappa = v$. We need the following two properties to be fulfilled:

Integrity. For every $\kappa \in \mathcal{K}$ and $e \in \mathcal{E}$, $(e)_\kappa \in \mathcal{E}$.

Commutativity. For every $\kappa, \kappa' \in \mathcal{K}$, $((v)_\kappa)_{\kappa'} = ((v)_{\kappa'})_\kappa$.

These properties hold when using any stream cipher, for which the ciphering operation consists in applying an *xor* between the datum and a stream of bits generated from the key. We need Integrity to ensure that players only send valid dissemination messages, regardless of the punishments. Commutativity is required for the scenarios when multiple punishments are applied simultaneously. More precisely, we need every player to be able to obtain disseminated events while not being punished. Moreover, we need every node to punish other nodes while being punished. This raises the following possibility. Consider nodes 1 to 4, and suppose that exactly both nodes 2 and 3 are being punished. Consider the following sequence of dissemination steps of event e : node 1 sends $e^1 = (e)_{\kappa_2}$ to node 2; node 2 sends $e^2 = (e^1)_{\kappa_3}$ to node 3; and node 3 sends an event e' to node 4. Since 4 is not being punished, we need to ensure that he is able to retrieve e from e' . If 3 were to send $((e)_{\kappa_2})_{\kappa_3}$, then this would only be possible if 4 knew that the event followed the path 1, 2, 3. Instead, 3 sends $e' = (e^2)_{\kappa_2}$. By Commutativity, it holds $e' = (e)_{\kappa_3}$. Since 4 knows κ_3 , he only needs to know that 3 is being punished in order to be able to retrieve e from e' .

With this in mind, a node being punished may receive any event e ciphered with κ_i , and possibly ciphered with κ_j for some $j \neq i$. In order for punishments to be effective, i must not be able to retrieve e when it has no access to κ_i , and has not received e in plain, regardless of what dissemination messages i may have received in the past. For this purpose, we need to ensure that: 1) i never receives κ_i , which is true by construction of our strategy; and 2) $\kappa_i \neq \kappa_j$ for every $j \in \mathcal{N} \setminus \{i\}$, ensuring that $((e)_{\kappa_i})_{\kappa_j} \neq e$ for any $e \in \mathcal{E}$. In addition, we need the following property to hold. Intuitively, before receiving any tuple with identifier id , i can guess the corresponding event $e \in \mathcal{E}^t$ with a small probability p . Non-disclosure ensures that i does not gain information that allows him to guess e with a probability significantly higher than p , provided that i only receives e ciphered with κ_i , and possibly with κ_j for some $j \in \mathcal{N} \setminus \{i\}$.

Formally, fix any $i \in \mathcal{N}$, $h_i \in \mathcal{H}_i$ from any stage t and $id \in \{1 \dots \nu\}$. Let $\mathcal{E}_i^t(h_i, id) \subseteq \mathcal{E}$ be the set of events such that for any $e \in \mathcal{E}_i^t(h_i, id)$ there exists $j \in \mathcal{N} \setminus \{i\}$ and dissemination round r from stage t preceding the observation of h_i such that j sends (id, e) to i in round r . We say that i has not received $e \in \mathcal{E}^t$ in h_i in plain iff for $e' = (e)_{\kappa_i}$:

$$\mathcal{E}_i^t(h_i, id) \subseteq \{e'\} \cup \{(e')_{\kappa_j} | j \in \mathcal{N} \setminus \{i\}\}.$$

Let $P(e|h_i)$ be the probability of i guessing e after the observation of h_i . We need the following property to hold:

Non-disclosure. For any $\epsilon > 0$, there exists a symmetric ciphering primitive such that, for any player $i \in \mathcal{N}$, private history $h_i \in \mathcal{H}_i$ from stage $t \in \mathbb{N}$, and event $e \in \mathcal{E}^t$ disseminated with identifier $id \in \{1 \dots \nu\}$, if i has not received e in plain in h_i , then for any $h'_i \in \mathcal{H}_i$ from stage t that fulfils $\mathcal{E}_i^t(h'_i, id) = \emptyset$ and $\mathcal{E}_i^t(h'_i, id') = \mathcal{E}_i^t(h_i, id')$ for every $id' \in \{1 \dots \nu\} \setminus \{id\}$, it holds:

$$P(e|h_i) \leq P(e|h'_i) + \epsilon.$$

These properties can be fulfilled by using a block ciphering algorithm with the CTR mode of operation [5, 11]. A careful selection of keys is also required. To simplify, we have the mediator selecting a new unique key per node in each stage, generated uniformly at random.

3.2 Pseudo-Random Number Generator

We assume the existence of a function SG_i^f per node i , defined as follows. Given a seed sd_i , this function returns a sequence of subsets, where $SG_i^f(sd_i, id) \subseteq \mathcal{N} \setminus \{i\}$ is the of f nodes to whom i must forward the event with identifier id . We need SG_i^f to fulfil the following requirements. Assuming that sd_i is chosen uniformly at random, the probability of $SG_i^f(sd_i, id)$ returning any subset S of f nodes is arbitrarily close to the probability p of i selecting S uniformly at random. Second, we need the stronger requirement of conditional independence. Namely, if we fix the subsets generated by SG_i^f for any identifier other than id , then the probability of $SG_i^f(sd_i, id)$ returning any subset is still arbitrarily close to p . This is formalised by the following property, where $P^{SG_i^f}(\vec{S} | \vec{S}_{-id})$ is the probability of $SG_i^f(sd_i, id)$ returning a sequence of subsets \vec{S} given that the seeds are selected uniformly at random among those that yield any sequence \vec{S}' such that $S_{id'} = S'_{id'}$ for every $id' \in \{1 \dots \nu\} \setminus \{id\}$, independently of the value of S'_{id} .

PRNG1. Fix any node $i \in \mathcal{N}$ and sequence of subsets $\vec{S} = (S_{id})_{id \in \{1 \dots \nu\}}$ with $S_{id} \subseteq \mathcal{N} \setminus \{i\}$ and $|S_{id}| = f$ for every $id \in \{1 \dots \nu\}$. For any constant $\xi > 0$, there exists a function SG_i^f such that for every $id \in \{1 \dots \nu\}$:

$$\left| P^{SG_i^f}(\vec{S} | \vec{S}_{-id}) - \frac{1}{\binom{n-1}{f}} \right| \leq \xi.$$

This property can be fulfilled by a PRNG function such as the one used in [2]. Additional details are included in Appendix B.

3.3 Strategy

It is useful to define $\vec{\sigma}^f$ using a state machine representation [17]. Each history $h_i \in \mathcal{H}_i$ is mapped into a private state of each player i . Transitions between states occur when players follow any action profile $\vec{a} \in \mathcal{A}$. Given any $h_i \in \mathcal{H}_i$, $\sigma_i^f(\cdot | h_i)$ suggests a probability distribution over the set of actions available to i after the observation of h_i . i may or may not follow this suggestion, such that a state may be reached where i knows he has deviated in the past. The only exception is acquiescent players, who never deviate. We now specify the state, transition rules, and strategy for every $i \in \mathcal{N}$, including the source.

3.3.1 State

Let $rn \in \{1 \dots \tau\}$ be the round number. Node i keeps for each node j a variable $\Theta_i(j) \in \{Good, Bad\}$, where $\Theta_i(j) = Good$ iff i has observed only valid actions from j . In addition, i keeps a set $Miss_i$ of sequences of identifiers that contain identifiers that were not forwarded appropriately according to SG_i^f . For each node j , i keeps two sets $RE_i(j)$ and $SE_i(j)$ of tuples (id, r) , representing events with identifier id received from and sent to j in round r , respectively. Finally, PE_i contains tuples (id, e) with an identifier id and an event e to be forwarded. An important aspect is that this state is finite, implying that memory is bounded.

3.3.2 Transition Rules

Algorithm 1 contains the pseudo-code. Every node initialises $\Theta_i(j) = Good$ at the beginning of each stage and sets $\Theta_i(j) = Bad$ when j does not follow a valid action (Lines 12 and 17). Valid actions are enumerated as follows. In monitoring rounds, each message has a fixed size. In the first monitoring round, node i must send to the mediator for each $j \neq i$ a message the size of an accusation against j . In any even monitoring round, for each node $j \neq i$ such that the mediator requested the corresponding sequence of identifiers, i must send two tuples per identifier id from the requested sequence. If sufficient information is not available, then i must send padding. In a dissemination round, i is allowed to send any set of tuples (id, e) , but only one tuple per identifier. Moreover, the identifier must not have expired and must have already been introduced. More precisely, define $age(id, r) = r + 1 - (id + \tau^M)$. Event with identifier id is introduced in round r such that $age(id, r) = 1$, i.e., $r = \tau^M + id$. A tuple containing id and sent in round r is valid iff $age(id, r) \in \{1 \dots \rho\}$. This restriction ensures that in any round every player has to forward at most ρ tuples to f nodes. To simplify, we assume that every node has sufficient bandwidth to send ρf tuples in a single round.

Node i registers in $RE_i(j)$ and $SE_i(j)$ for each $id \in \{1 \dots \nu\}$ the round number when j first sent to and received from i a tuple containing id , in a valid dissemination message (Lines 19 and 23). When i receives id for the first time and id has not expired ($age(id, rn) \leq \rho$), i selects a tuple (id, e) to be forwarded, sent by some node j chosen according to some deterministic rule such as the node with smallest identifier (Lines 25-30). Not every deterministic rule is allowed, since we need to ensure that in any two scenarios 1 and 2 where the set of nodes S sending id to i in scenario 1 is a subset of scenario 2, and $j \in S$ is selected in 2, then j is also selected in 1. An alternative valid rule would be the node with the largest identifier. If j is being punished, then i forwards $(id, (e)_{\kappa_j})$ instead (Line 29). Let $seq(id)$ denote the sequence to which identifier id belongs to. i adds $seq(id)$ to $Miss_i$ whenever i knows that, if the mediator requests reports relative to $seq(id)$, then i will be punished (Line 21). This occurs exactly when the reports relative to id are inconsistent: i sends a tuple containing id to $j \notin SG_i(sd_i, id)$ (Line 33), i fails to send id to $j \in SG_i(sd_i, id)$ immediately after the first reception of id (Line 38), or i sends id prior to receiving it (Line 40). Notice that we update RE_i after updating $Miss_i$, such that identifiers first received in the present round do not count as being inconsistent.

3.3.3 Strategy Definition

The pseudo-code is included in Algorithm 2. Player i stops

Algorithm 1 Transition Rules

```

1: After first monitoring round
2:    $rn \leftarrow 1$ 
3:    $PE_i, Miss_i \leftarrow \emptyset$ 
4:   for all  $j \in \mathcal{N}$  do
5:      $RE_i(j), SE_i(j) \leftarrow \emptyset$ 
6:      $\Theta_i(j) \leftarrow Good$ 
7:   Set  $\Theta_i(i) = Bad$  if  $i$  sent invalid message
8: End

9: After any other monitoring round
10:   $rn \leftarrow rn + 1$ 
11:  for all  $j \in \mathcal{N}$  do
12:    Update  $\Theta_i(j)$ 
13: End

14: After dissemination round
15:   $rn \leftarrow rn + 1$ 
16:  for all  $j \in \mathcal{N}$  do
17:    Update  $\Theta_i(j)$ 
18:    for all valid  $id$  sent by  $i$  to  $j$  for the first time do
19:      Add  $(id, rn)$  to  $SE_i(j)$ 
20:    for all  $id \in \{1 \dots \nu\}$ : inconsistent( $id, SE_i, RE_i$ ) do
21:      Add  $seq(id)$  to  $Miss_i$ 
22:    for all valid  $id$  sent by  $j$  to  $i$  for the first time do
23:      Add  $(id, rn)$  to  $RE_i(j)$ 
24:    for all  $id \in \{1 \dots \nu\}$  received for the first time do
25:      if  $age(id, rn) \leq \rho$  then
26:         $j \leftarrow$  node with smallest identifier to send  $id$ 
27:         $e \leftarrow$  event such that  $j$  sent  $(id, e)$ 
28:        if  $i$  received accusation against  $j$  then
29:           $e \leftarrow (e)_{\kappa_j}$ 
30:        Add  $(id, e)$  to  $PE_i$ 
31: End

32: Predicate inconsistent( $id, S, R$ )
33:   if  $\exists_{j \notin SG_i^f(sd_i, id)} \exists_r (id, r) \in S(j)$  then
34:     return True
35:   else if  $\exists_{j, r} (id, r) \in R(j)$  then
36:      $r \leftarrow$  minimum round when  $i$  received  $id$  in  $R(j)$ 
37:     if  $age(id, r) \in \{1 \dots \rho - 1\}$  then
38:       if  $\exists_{l \in SG_i^f(sd_i, id)} (id, r + 1) \notin S(j)$  then
39:         return True
40:   else if  $\exists_{j, r} (id, r) \in S(j)$  then
41:     return True
42:   else
43:     return False
44: EndPredicate

```

sending messages once $\Theta_i(i) = Bad$, since a punishment in the next stage is inevitable. In the first monitoring round, i sends an accusation against $j \neq i$ iff $\Theta_i(j) = Bad$ at the end of the previous stage (Line 4). In an even monitoring round, the mediator notifies i for each node $j \neq i$ whether i should report on the sequence S of identifiers corresponding to the current round, with probability p^* (Line 9). When i receives this notification, i sends for each identifier $id \in S$ any tuples $(id, r) \in RE_i(j)$ and $(id, r') \in SE_i(j)$ (Lines 12-16). In the last monitoring round, the mediator notifies every node $l \neq j$ that j must be punished when some node $k \neq j$ has sent an accusation against j , or he detects an inconsistency regarding some $id \in \{1 \dots \nu\}$, according to the reports RE_j and SE_j regarding events received and sent by j , respectively (Lines 19-27). The mediator also sends sd_i to i and κ_i to every $j \neq i$, ensuring that κ_i is unique. In a dissemination round, i sends any tuple $(id, e) \in PE_i$ to every $j \in SG_i(sd_i, id)$, ciphering e with κ_j if j is being punished (Lines 30-34). Notice that i only sends a tuple

Algorithm 2 Strategy

```

1: Upon first monitoring round of stage  $t > 1$ 
2:   for all  $j \in \mathcal{N}$  do
3:     if  $\Theta_i(j) = \text{Bad}$  then
4:       Send accusation against  $j$  to mediator
5:     Send padding
6: End

7: Upon even monitoring round  $rn < \tau^M$  |  $i$  is the mediator
8:   for all  $j \in \mathcal{N}$  do
9:     With probability  $p^*$ , send  $(j, \text{Monitor})$  to every  $l \neq j$ 
10: End

11: Upon odd monitoring round |  $1 < rn < \tau^M$  and  $\Theta_i(i) = \text{Good}$ 
12:   for all  $j \in \mathcal{N} \setminus \{i\}$ : mediator sent  $(j, \text{Monitor})$  do
13:      $S_j \leftarrow$  current sequence of identifiers
14:     Send all tuples  $(id, r) \in RE_i(j)$  with  $id \in S_j$ 
15:     Send all tuples  $(id, r) \in SE_i(j)$  with  $id \in S_j$ 
16:     Send padding
17: End

18: Upon last monitoring round  $\tau^M$  |  $i$  is mediator
19:   for all  $j \in \mathcal{N}$  do
20:     if  $\exists_{l \neq j}$   $l$  sent accusation against  $j$  then
21:       Send accusation against  $j$  to every  $o \in \mathcal{N}$ 
22:     else if  $\exists_{id}$  inconsistent  $(id, SE_j, RE_j)$  then
23:       Send accusation against  $j$  to every  $l \in \mathcal{N}$ 
24:      $sd_j \leftarrow$  random value
25:     Send  $sd_j$  to  $j$ 
26:      $\kappa_j \leftarrow$  unique random key
27:     Send  $\kappa_j$  to every  $l \neq j$ 
28: End

29: Upon dissemination round |  $\Theta_i(i) = \text{Good}$ 
30:   for all  $(id, e) \in PE_i : seq(id) \notin Miss_i$  do
31:     for all  $j \in SG_i(sd_i, id)$  do
32:       if  $i$  has accusation against  $j$  then
33:          $e \leftarrow (e)_{\kappa_j}$ 
34:       Send  $(id, e)$  to  $j$ 
35: End

36: Upon dissemination of event  $e \in \mathcal{E}^t$  |  $i$  is the source
37:    $id \leftarrow rn - \tau^M + 1$ 
38:   Add  $(id, e)$  to  $PE_i$ 
39: End

```

(id, e) if $seq(id) \notin Miss_i$. This is because once i fails to send some identifier from sequence $seq(id)$, sending id as specified by SG_i does not affect the probability of being punished in the next stage. Thus, it is optimal to drop every tuple with an identifier from that sequence. Later, we specify how to define p^* such that for every identifier id with $seq(id) \notin Miss_i$ it is still optimal for i to forward id .

3.4 Realised Utility

We now define $u_i(z)$ for every outcome $z \in \mathcal{Z}^t$ of any stage t . We consider that every player i incurs a fixed cost α per bit sent in any message. In addition, i obtains a benefit β per event $e \in \mathcal{E}^t$ received by i during stage t . We denote the set of received tuples by $rec_i(z)$, where for every $(id, e) \in rec_i(z)$ we have $e \in \mathcal{E}^t$. Two issues arise when trying to define $rec_i(z)$: 1) the exact definition of i receiving (id, e) and 2) the reception of ciphered events.

Regarding the first issue, since we are considering reliable communication channels, i receives (id, e) iff some node $j \neq i$ sends (id, e) to i . Thus, we may consider that i obtains a benefit β in this case. However, given our definition of strategies, it is possible that i sends (id, e) to some node before receiving it in the first place, and then this tuple loops

back to i . Such behaviour increases the expected utility of i . Since i is unlikely to guess e beforehand, we consider that i values e iff i knows that it could not have been introduced by i . More precisely, i receives a benefit β per disseminated tuple (id, e) with $e \in \mathcal{E}^t$ iff some $j \neq i$ sends (id, e) to i in round r of stage t , and i has not sent (id, e) to any node in any round $r' < r$.

Regarding the second issue, it is possible that i never receives some $e \in \mathcal{E}^t$ in plain, but instead ciphered with the key κ_j of some node $j \neq i$. We define the strategy in a way that, if $e' = (e)_{\kappa_j}$ and i receives e' , then it is because j is being punished, j has sent e' to i , and i knows that j is being punished. Hence, i is able to compute $(e')_{\kappa_j}$, in which case we say that i can retrieve e from e' . Naturally, i can retrieve e whenever i receives e in plain. In addition, it may be possible that a history is reached where i receives e ciphered with multiple different keys and i is still able to retrieve e . To generalise this intuition, we say that i is able to retrieve e from e' whenever i can perform some computation over e' , given the history of interactions with other nodes, in order to obtain e . i can retrieve e from e' when $e = e'$. Also, if i can retrieve e from e' , i receives $e'' = (e')_{\kappa_j}$ from j , and i knows that j is being punished with key κ_j , then i can retrieve e from e'' . By Non-disclosure, if some j sends (id, e') to i and $e' = (e)_{\kappa_i}$ or $e' = ((e)_{\kappa_i})_{\kappa_j}$, then i cannot retrieve e from e' . We do not make any further assumptions regarding when i can retrieve e from e' .

With this in mind, we consider that, for any $e \in \mathcal{E}^t$ disseminated by the source with identifier id , we have $(id, e) \in rec_i(z)$ iff there exists a round r from stage t and node $j \in \mathcal{N} \setminus \{i\}$ such that j sends (id, e') to i in round r , i can retrieve e from e' , and i has not sent (id, e') to any node in any round $r' < r$ from stage t . This leads to the following definition of realised utility. Recall that we normalise the total benefits and costs to the average per disseminated event, by dividing it by ν . Let $|z_i^{t,r}(j)|$ represent the size of the message sent by i to $j \in \mathcal{N}$ in round r of stage t :

$$u_i(z) = \frac{1}{\nu} \left(\beta \cdot |rec_i(z)| - \alpha \cdot \sum_{r=1}^{\tau} \sum_{j \in \mathcal{N} \setminus \{i\}} |z_i^{t,r}(j)| \right). \quad (3)$$

4. ANALYSIS

The analysis is divided into three parts. First, we show correctness properties of $\bar{\sigma}^f$ for any $f > 0$. Then, we compare the utility of following $\bar{\sigma}^f$ with that of deviating. We conclude with the proof of the main result.

We use the notation v^h and v^{h_i} to denote the value of state variable v after any histories h and h_i are realised, respectively. In this context, we say that a history h' succeeds h if h' is reached with positive probability when players follow $\bar{\sigma}^f$ after h ; h' immediately succeeds h if h' is reached one round after h . Similarly, h' precedes h if h' is a starting subsequence of h ; h' immediately precedes h if $rn^h = rn^{h'} + 1$.

The proof relies the following facts regarding any node i and history h from stage t : i) for any $z \in \mathcal{Z}^t$ succeeding h , $\Theta_i^z(i) = \Theta_i^h(i)$ and $Miss_i^z = Miss_i^h$; ii) $\Theta_i^h(i) = \text{Good}$ if and only if, for every $j \in \mathcal{N}$ and $id \in \{1 \dots \nu\}$, $\Theta_j^h(i) = \text{Good}$; and iii) for any sequence of identifiers s , we have $s \in Miss_i^h$ if and only if there exists $id \in \{1 \dots \nu\}$ such that $s = seq(id)$ and the reports relative to i are inconsistent with regard to id , such that if every player follows $\bar{\sigma}^f$ in the future, then i is punished if and only if the mediator requests the identifiers

from a sequence in $Miss_i^h$.

The analysis relies on the following proposition, which follows by construction of the strategy.

PROPOSITION 4.1. *For any history $h \in \mathcal{H}$ from stage $t \in \mathbb{N}$ and node $i \in \mathcal{N}$, the following hold: 1) for any $z \in \mathcal{Z}^t$ succeeding h , $\Theta_i^z(i) = \Theta_i^h(i)$ and $Miss_i^z = Miss_i^h$; 2) $\Theta_i^h(i) = \text{Good}$ iff, for every $j \in \mathcal{N} \setminus \{i\}$, $\Theta_j^h(i) = \text{Good}$; and 3) for any identifier $id \in \{1 \dots \nu\}$, we have $id \in Miss_i^h$ iff there exists $id' \in seq(id)$ such that the reports relative to i are inconsistent with regard to id' .*

PROOF. Fix t , h , and i . Fix z succeeding h . i only changes $\Theta_i(i)$ from Good to Bad . Thus, if $\Theta_i^h(i) = \text{Bad}$, then $\Theta_i^z(i) = \text{Bad}$. Otherwise, i only changes $\Theta_i^h(i)$ after h if i sends an invalid message, which never occurs while i follows σ_i^f . Therefore, $\Theta_i^h(i) = \Theta_i^z(i)$. By construction, while following σ_i^f , i forwards every tuple with identifier id and $seq(id) \notin Miss_i^h$, first received in round r , to exactly all the nodes from $SG_i^f(sd_i, id)$ in round $r + 1$. Thus, i never adds any new sequence to $Miss_i^h$ while following σ_i^f , implying that $Miss_i^h = Miss_i^z$. This proves 1). Now, recall that every $j \in \mathcal{N}$ initialises $\Theta_j(i) = \text{Good}$ at the beginning of stage t . If $\Theta_i^h(i) = \text{Bad}$, then i sent an invalid message to some $j \neq i$, implying that $\Theta_j^h(i) = \text{Bad}$. Otherwise, no j updated $\Theta_j(i)$ to Bad . This proves 2). Finally, since i initialises $Miss_i = \emptyset$ in stage t , we have $id \in Miss_i^h$ iff there exists $id' \in seq(id)$ such that $inconsistent(id', SE_i^h, RE_i^h)$ is true. By construction, this holds iff the reports relative to i are inconsistent with regard to id' . This proves 3). \square

4.1 Correctness

We show two sets of properties regarding monitoring and dissemination. Monitoring properties characterise the probability of any node $i \in \mathcal{N}$ being punished in the present or future stages as a function of the current state. Dissemination properties quantify the probability of any node i receiving events, already disseminated or disseminated only in the future, as a function of the current state and the present action \bar{a}_i of i . This allows us to compute the expected utility of any player i for each possible action he takes in the present, and this way prove that our strategies are SE by applying the One-deviation Property.

Lemma 4.2 enumerates monitoring Properties M1-M4 valid for any player i and history h , assuming that every player follows $\bar{\sigma}^f$ after h . M1-M3 state that the probability of i being punished in the next stage is a function of $\Theta_i^h(i)$ and $Miss_i^h$, while i is never punished in future stages other than the next. In addition, Property M4 shows that i cannot influence the punishments being applied to him in the present stage by only deviating in the current round.

LEMMA 4.2. *For any history $h \in \mathcal{H}$ from stage $t \in \mathbb{N}$ and player $i \in \mathcal{N}$, the following properties hold:*

M1. *If $\Theta_i^h(i) = \text{Good}$, then the mediator accuses i in stage $t + 1$ with probability $P(|Miss_i^h|) = 1 - (1 - p^*)^{|Miss_i^h|}$.*

M2. *If $\Theta_i^h(i) = \text{Bad}$, then the mediator accuses i in stage $t + 1$ with probability 1.*

M3. *The mediator never accuses i in any stage $t' > t + 1$.*

M4. *For any two actions $\bar{a}_i, \bar{a}'_i \in \mathcal{A}_i$ and round $r > rn^h$ from stage t , i is punished in round r either by every node or by no node, the probability of i being punished is the same*

after \bar{a}_i^ and \bar{a}'_i , and every node $j \in \mathcal{N} \setminus \{i\}$ gets the same key κ_i .*

PROOF. **M1.** Fix any $z \in \mathcal{Z}^t$ succeeding h . By Proposition 4.1, we have $\Theta_i^z(i) = \text{Good}$ and $Miss_i^z = Miss_i^h$, and $\Theta_j^z(i) = \text{Good}$ for every $j \neq i$. Thus, no node j sends an accusation against i to the mediator in the first round of stage $t + 1$. Moreover, since nodes never remove entries from RE and SE , for every sequence $S \in Miss_i^h$, reports relative to i after z are inconsistent regarding some $id \in S$. If the mediator selects S in the corresponding even monitoring round of stage $t + 1$, then every node sends his reports relative to i and regarding every event from S , including id . Thus, the mediator detects the inconsistency regarding id and emits an accusation against i in the last monitoring round of stage $t + 1$. This occurs with probability $P(|Miss_i^z(i)|) = P(|Miss_i^h(i)|)$.

M2. Fix any $z \in \mathcal{Z}^t$ succeeding h . By Proposition 4.1, it holds $\Theta_j^z(i) = \text{Bad}$ for some $j \neq i$. Hence, j sends an accusation against i to the mediator in the first monitoring round of stage $t + 1$, and the mediator accuses i in the last monitoring round of stage $t + 1$.

M3. Let $t' \geq t + 1$ be any stage number and fix any $z \in \mathcal{Z}^{t'-1}$ succeeding h . Since i follows $\bar{\sigma}^f$ in the first monitoring round of stage t' , we have $\Theta_i^{h'}(i) = \text{Good}$ and $Miss_i^{h'} = \emptyset$ for h' immediately succeeding z . By M1, it follows that the mediator accuses i in stage $t' + 1$ with probability:

$$P(|Miss_i^{h'}|) = P(0) = 0.$$

M4. Regarding the first and last statements, recall that the mediator is trusted. Hence, we are focusing on any history h where he has not deviated in the past. Therefore, whether h precedes the last monitoring round, the mediator always sends the same key κ_i to every $j \in \mathcal{N} \setminus \{i\}$, and either sends an accusation to every node or to none. Regarding the second statement, the mediator ignores any information sent by i when determining whether i should be punished. If $h \in \mathcal{Z}^{t-1}$, then the probability of i being accused by the mediator depends solely on $\Theta_j(i)$, $rec_j(i)$, and $SE_j(i)$ for every $j \in \mathcal{N} \setminus \{i\}$, and on the probability p^* of the mediator selecting each sequence of identifiers. Otherwise, the probability of i being punished depends on the same information except $\Theta_j(i)$, and depends on what j already sent to the mediator in previous monitoring rounds of stage t . If h follows the last monitoring round, then either every node $j \in \mathcal{N} \setminus \{i\}$ already has an accusation against i in h , or none has, regardless of the present action. \square

Lemma 4.3 enumerates properties D1-D5. The complete proof is in Appendix A. Given any history h from stage t , event e , and node $i \in \mathcal{N}$, let $g^h(e, i) = (e)_{\kappa_i}$ if the mediator triggers a punishment of i in stage t with key κ_i , or $g^h(e, i) = e$ otherwise. By M4, we have that for any $j \neq i$, if $(id, e') \in PE_j$ and j forwards id to i , then j sends $g^h(e', i)$. Conversely, if j receives (id, e') for the first time from i , then j adds $(id, g^h(e', i))$ to PE_j . For any history h and two actions \bar{a}_i^* and \bar{a}'_i , we use the notation $v_i^*(i)$ and $v_i'(i)$ to represent the value of any state variable v_i resulting from i following \bar{a}_i^* and \bar{a}'_i after h , respectively. For some j and message $a_i(j)$, we consider that $id \in a_i(j)$ for some j if there exists a tuple $(id, e) \in a_i(j)$.

Properties D1-D3 refer to events introduced only in the future. Namely, Property D1 states that, regardless of the punishments being applied, a player i obtains a disseminated

tuple (id, e) iff i is not punished and receives $(id, g^h(e, j))$ from some j . D2 and D3 indicate that i cannot influence the probability of receiving any event not yet introduced. Properties D4-D5 refer to events being disseminated. More precisely, D4 states that if i follows two alternative actions in which i sends a given identifier id already disseminated to the same set of nodes, then i receives a tuple containing id after both actions with the same probability. D5 states that if the source previously disseminated a tuple (id, e) and i follows two actions \vec{a}_i^* and \vec{a}_i' where i does not send more tuples containing id in \vec{a}_i^* than in \vec{a}_i' to any node, then the probability of i retrieving e after following \vec{a}_i^* is at least as high as after following \vec{a}_i' . The complete proofs are in Appendix A.

LEMMA 4.3. *Fix any history $h \in \mathcal{H}$ from stage $t \in \mathbb{N}$, player $i \in \mathcal{N}$, and identifier $id \in \{1 \dots \nu\}$. The following properties hold:*

D1. *If the source introduces (id, e) after the realisation of h , then for all history $h^* \in \mathcal{H}$ from stage t succeeding h , player $j \in \mathcal{N}$, and $(id, e') \in PE_j^h$, it holds $e' = g^h(e, j)$.*

D2. *For every $t' > t$ and any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i$, the probability of i receiving some tuple containing identifier id in stage t' after i follows \vec{a}_i^* is the same as after \vec{a}_i' .*

D3. *If id is introduced in stage t after h is realised, then for any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i$, the probability of i receiving some tuple containing identifier id in stage t after i follows \vec{a}_i^* is the same as after i follows \vec{a}_i' .*

D4. *Fix any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i$ that fulfil: i) $\Theta_i^*(i) = \Theta_i'(i)$ and $Miss_i^* = Miss_i'$; and ii) for any $j \in \mathcal{N}$, $id \in a_i^*(j)$ iff $id \in a_i'(j)$. If id is disseminated in stage t before h is realised, then i receives some tuple containing id after following \vec{a}_i^* iff the same holds after following \vec{a}_i' .*

D5. *Let $e \in \mathcal{E}^t$ be the event disseminated with identifier id . Fix any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i^h$ that fulfil: i) $\Theta_i^*(i) = \Theta_i'(i)$ and $Miss_i^* = Miss_i'$; and ii) for each $j \in \mathcal{N}$, $id \in a_i^*(j)$ only if $id \in a_i'(j)$. For every outcomes $z^*, z' \in \mathcal{Z}^t$ succeeding \vec{a}_i^* and \vec{a}_i' , respectively, if id is introduced in stage t prior to h being realised and $(id, e) \in rec_i(z')$, then $(id, e) \in rec_i(z^*)$.*

PROOF. (Sketch) **D1.** The source forwards $(id, g^h(e, i))$ or no tuple containing id to each i . Inductively, if we suppose that D1 holds for round r , then any j sending (id, e') to i in round $r + 1$ has $e' = g^h(g^h(e, j), i)$, while i adds (id, e'') to PE_i , where $e'' = g^h(e, j)$. By Commutativity, $e'' = g^h(e, i)$.

D2. We set $\Theta_j(j) = Good$ and $Miss_j = \emptyset$ after first round of stage t' , and by Proposition 4.1 these values are preserved throughout stage t' . Thus, every node j forwards id immediately after its first reception to l iff $l \in SG_j^f(sd_j, id)$. Since the probability distribution over seeds is the same after both actions, then so is the probability of i receiving id .

D3. This probability depends only on the value $\Theta_j(j)$, $Miss_j$, and sd_j for every $j \in \mathcal{N} \setminus \{i\}$: every node follows $\vec{\sigma}^f$ after h , during the dissemination of id ; hence, i receives id iff there exists a path of nodes terminating in i such that every j in that path forwards id to the next node l , which occurs when $\Theta_j(j) = Good$, $seq(id) \notin Miss_j$, and $l \in SG_j^f(sd_j, id)$. By Proposition 4.1, the values $\Theta_j(j)$ and $Miss_j$ are never updated after h , regardless of what i does in the present. If h follows the last monitoring round, then sd_j is already fixed since mediator is trusted, such that i either receives id after h or not, independently of the current action. Other-

wise, the probability distribution guiding the choice of sd_j is independent of the actions of i .

D4. We use induction to show the result, starting from the base case that any node j receives id for the first time immediately after h with the same probability, regardless of the action of i . Also, by Proposition 4.1, $\Theta_j(j)$ and $Miss_j$ are preserved after h for every $j \in \mathcal{N} \setminus \{i\}$, regardless of the current action of i . Thus, j either forwards id after h and after first receiving id to any node l after \vec{a}_i^* iff j does the same after \vec{a}_i' . This fact implies the induction step.

D5. We prove two sub-properties. First, we show that, by not sending more identifiers in \vec{a}_i^* , the number of identifiers being disseminated at each point in time after \vec{a}_i^* is a subset of that after \vec{a}_i' , in a similar fashion to D4. Using this property, we show that, if $(id, e) \in rec_i(z')$ and i did not receive (id, e) in h , then it must hold $(id, e) \in rec_i(z^*)$. This follows from three facts. First, i can never forward (id, e') in \vec{a}_i' before receiving it while being able retrieve e from e' , since otherwise we would have by definition $(id, e) \notin rec_i(z')$. Second, if i forwards (id, e') in \vec{a}_i' and this tuple loops back to i , then, despite nodes successively applying ciphers over e' , i always receives (id, e') back. Thus, we avoid the scenario where i receives $e' = (e)_{\kappa_i}$, forwards it causing some node to decipher e , and expects (id, e) to return to i without some node ciphering e again. The third fact is the following. Suppose that i believes that some tuple (id, e') is held by some node j , and that this tuple is forwarded all the way to i after i follows \vec{a}_i' . Suppose also that i can retrieve e from the event resulting from the successive ciphering operations applied over e' . Then, the exact same tuple reaches i when i follows \vec{a}_i^* . Here, the fact that nodes deterministically forward the event received from the node with smallest identifier plays a key role. More precisely, by the first property, the set of identifiers forwarded by any node after \vec{a}_i^* is a subset of the same set after \vec{a}_i' . Thus, whenever some j forwards (id, e') to l after \vec{a}_i' and l prepares to forward e' because j has the smallest identifier, then j also forwards (id, e') to l and has the smallest identifier after \vec{a}_i^* such that l prepares to forward e' . \square

Lemma 4.4 enumerates dissemination Properties D6 and D7, which quantify the probability of node i receiving each disseminated event. D6 states that this probability is arbitrarily close to $q(f)$ for events disseminated in future stages, which is true by PRNG1. D7 states that for any event with identifier id introduced in future rounds of the present stage this probability is at most $q(f) + \xi$ for an arbitrarily small ξ . This follows from the fact that i learns for each player $j \neq i$ at most the value $SG_j^f(sd_j, id')$ for every id' previously introduced, but i does not learn these values for id . Hence, any seed yielding any subset $SG_j^f(sd_j, id)$ is possible and equally likely to have been chosen by the mediator, independently of the subsets generated for other events. By Consistency of $(\vec{\sigma}^f, \mu^f)$, i believes this to be true. By PRNG1, it follows that the probability of any node receiving id is at most arbitrarily close to $q(f)$, although it may be lower due to the existence of players that drop id .

LEMMA 4.4. *Fix any player $i \in \mathcal{N}$, stage $t \in \mathbb{N}$, private history $h_i \in \mathcal{H}_i$ from stage t , and identifier $id \in \{1 \dots \nu\}$. The following properties hold:*

D6. *For every constant $\xi > 0$ and stage $t' > t$, there exists a function SG_i such that i receives some tuple con-*

taining identifier id in stage t' with probability q' fulfilling $|q' - q(f)| < \xi$.

D7. If id is introduced in stage t after h_i is observed, then for every constant $\xi > 0$ there exists a function SG_i such that i believes that will receive some tuple containing identifier id in stage t with probability at most $q(f) + \xi$.

4.2 Utility Analysis

In line with the One-deviation property, the goal of this section is to compute the difference between the expected utilities $\mathbb{E}^{\vec{\sigma}^f, \mu}(u_i|h_i, \vec{a}_i) - \mathbb{E}^{\vec{\sigma}^f, \mu}(u_i|h_i, \vec{a}'_i)$ of player i following two alternative actions \vec{a}_i and \vec{a}'_i after any private history h_i . We divide the difference between expected utilities calculated from stage t into three parts: *long-term* utilities, referring to any stage $t' > t + 1$; *medium-term* utilities relative to stage $t + 1$; and *short-term* utilities relative to stage t . We will denote by γ the cost of sending a tuple, i.e., $\gamma = \alpha(c + \lceil \log(\nu) \rceil)$, where c is the size of an event

4.2.1 Long-term

Lemma 4.5 shows that the expected utility of any stage $t' > t + 1$ is fixed, regardless of the present action in stage t . The reason for this is that by our definition of strategy, regardless of what player i does in stage t , he sends every requested message in stage t' . Also, since he does not deviate in any future stage, he is never punished in stage t' , receiving all events with a fixed probability.

LEMMA 4.5. Fix any player $i \in \mathcal{N}$, private history $h_i \in \mathcal{H}_i$ from stage $t \in \mathbb{N}$, and any two actions $\vec{a}_i^*, \vec{a}'_i \in \mathcal{A}_i$. For every $t' > t + 1$, we have:

$$\mathbb{E}^{\vec{\sigma}^f, \mu, t'}(u_i|h_i, \vec{a}_i^*) = \mathbb{E}^{\vec{\sigma}^f, \mu, t'}(u_i|h_i, \vec{a}'_i).$$

PROOF. By construction, i sets $\Theta_i(i) = \text{Good}$ and $\text{Miss}_i = \emptyset$ at the beginning of stage t' , regardless of whether i follows \vec{a}_i^* or \vec{a}'_i . By Proposition 4.1, these variables are never updated during stage t' . Thus, i always sends every requested monitoring message, implying that expected monitoring costs for stage t' are the same whether i follows \vec{a}_i^* or \vec{a}'_i . By D2, i receives a tuple with each identifier $id \in \{1 \dots \nu\}$ with the same probability after both actions. Since $\Theta_i(i) = \text{Good}$ and $\text{seq}(id) \notin \text{Miss}_i = \emptyset$ after any round when i receives id for the first time, i forwards a tuple containing id to exactly f nodes with the same probability after both actions. This implies that expected costs of disseminating events are the same. Finally, by M3 and M4, in any dissemination round r of stage t' , no node $j \in \mathcal{N} \setminus \{i\}$ has an accusation of i . By D1, if i receives (id, e') from j , then $e' = g^h(e, j)$ such that i is capable of retrieving e from e' . Also, i never forwards id before receiving a tuple with this identifier. Therefore, $(id, e) \in \text{rec}_i(z)$ for any $z \in \mathcal{Z}^{t'}$ reached with positive probability such that i a tuple containing id in z . By D2, i receives a tuple containing id with the same probability after following \vec{a}_i^* and \vec{a}'_i . Therefore, the expected benefits are the same. \square

4.2.2 Medium-term

Lemma 4.9 analyses the difference between the expected utility of player i following an action \vec{a}_i^* prescribed by the strategy for dissemination rounds and that of following an arbitrary action \vec{a}'_i . We consider three cases, which cover all the scenarios of interest to show that $\vec{\sigma}^f$ is an equilibrium.

Namely, we have $\Theta_i^*(i) = \Theta_i^{h_i}$ and $\text{Miss}_i^* = \text{Miss}_i^{h_i}$ for any observed $h_i \in \mathcal{H}_i$ such that i follows \vec{a}_i^* after h_i . This is because in \vec{a}_i^* player i forwards every event as specified by the strategy and does not send invalid messages. By following \vec{a}'_i , i may not update $\Theta_i(i)$ or Miss_i by following valid actions only and forwarding every event as expected, may update $\Theta_i(i)$ from *Good* to *Bad* after sending an invalid message, or add multiple sequences to Miss_i when not forwarding some events as expected. However, i may never update $\Theta_i(i)$ from *Bad* to *Good* nor remove an element from Miss_i .

LEMMA 4.6. Fix any player $i \in \mathcal{N}$, private history $h_i \in \mathcal{H}$ from stage $t \in \mathbb{N}$, and any two actions $\vec{a}_i^*, \vec{a}'_i \in \mathcal{A}_i$. The following holds regarding the value $\Delta^{t+1} = \mathbb{E}^{\vec{\sigma}^f, t+1}(u_i|h, \vec{a}_i^*) - \mathbb{E}^{\vec{\sigma}^f, t+1}(u_i|h, \vec{a}'_i)$:

1. If $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$, then $\Delta^{t+1} \geq 0$.
2. If $\Theta_i^*(i) = \text{Good}$ and $\Theta_i'(i) = \text{Bad}$, then for any $\xi > 0$ there exists a function SG_i^f such that:

$$\Delta^{t+1} \geq (q(f) - \xi)\beta(1 - P(|\text{Miss}_i^*|)). \quad (4)$$

3. If $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$, then for any $\xi > 0$ there exists a function SG_i^f such that:

$$\Delta^{t+1} \geq (q(f) - \xi)\beta(P(|\text{Miss}_i'|) - P(|\text{Miss}_i^*|)). \quad (5)$$

PROOF. Like in the previous lemma, i initialises $\Theta_i(i) = \text{Good}$ and $\text{Miss}_i = \emptyset$ in stage $t + 1$, while never updating these variables during this stage as stated by Proposition 4.1. Thus, expected monitoring costs are the same whether i follows \vec{a}_i^* or \vec{a}'_i . By D2, i receives a tuple for each $id \in \{1 \dots \nu\}$ with a fixed probability independent of the current action, and forwards id to f nodes. Thus, expected dissemination costs are also the same after both actions for any case. Now, we analyse the expected benefits individually for each case.

Case 1. This case is very similar to Lemma 4.5. By M2, i is punished with probability 1 after both actions. By construction, if any $j \neq i$ sends (id, e') to i , then $(id, e') \in PE_j$ where $e'' = g^h(e', i) = (e')_{\kappa_i}$. By D1, $e'' = g^h(e, j)$, given that the source introduces (id, e) . Thus, for every (id, e') sent to i , either we have $e' = (e)_{\kappa_i}$ or $((e)_{\kappa_i})_{\kappa_j}$ for some $j \neq i$. By non-disclosure, for any $z \in \mathcal{Z}^{t+1}$ reached with positive probability after any action \vec{a}_i^* or \vec{a}'_i , $\text{rec}_i(z) = \emptyset$, such that expected benefits are 0.

Case 2. By M2, i is punished in stage $t + 1$ with probability 1 after following \vec{a}'_i . As in Case 1, by D1, the expected benefit is 0. On the other hand, by M1, i is punished in stage $t + 1$ after following \vec{a}_i^* with probability $P(|\text{Miss}_i^*|)$. When i is not being punished, by D1 and D6, i obtains a benefit β per disseminated event $e \in \mathcal{E}^{t+1}$ with probability at least $q(f) - \xi$ for an arbitrarily small ξ : i receives a tuple (id, e') from some j after any history h with probability at least $q(f) - \xi$, such that $e' = g^h(e, j)$ and i is able to retrieve e from e' ; thus, for at least a fraction $q(f) - \xi$ of the histories $z \in \mathcal{Z}^{t+1}$ reached with positive probability after \vec{a}_i^* or \vec{a}'_i , we have $(id, e) \in \text{rec}_i(z)$. Therefore, the expected benefits after i follows \vec{a}_i^* are at least $(q(f) - \xi)\beta\nu(1 - P(|\text{Miss}_i^*|))/\nu$, as we intended to prove.

Case 3. By D1, i retrieves each disseminated event iff i is not being punished. By M1, the probability of i being punished after following \vec{a}_i^* and \vec{a}'_i is $x^* = P(|\text{Miss}_i^*|)$ and $x' = P(|\text{Miss}_i'|)$, respectively. By D2 and D6, i receives each

event with some probability $q' \geq q(f) - \xi$ for an arbitrarily small $\xi > 0$, both after \vec{a}_i^* and \vec{a}_i' . This yields a difference between the expected benefits given as follows:

$$\begin{aligned} & (q'\beta\nu(1-x^*) - q'\beta\nu(1-x'))/\nu = \\ & = q'\beta(x' - x^*). \end{aligned}$$

This concludes the proof. \square

4.2.3 Short-term

We need to differentiate between the cases where the alternative actions are followed in dissemination and monitoring rounds, analysed in Lemmas 4.9 and 4.10, respectively. These lemmas use the following two auxiliary lemmas that analyse the expected utility relative to events disseminated in the present stage, after the current round.

Lemma 4.7 shows that the expected benefits of receiving events disseminated in future rounds are fixed, regardless of the present action.

LEMMA 4.7. *Fix any player $i \in \mathcal{N}$, history $h_i \in \mathcal{H}_i$ from stage $t \in \mathbb{N}$, $id \in \{1 \dots \nu\}$ introduced after h_i is observed, and any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i$. The expected benefits of receiving $e \in \mathcal{E}^t$ introduced by the source with identifier id after i follows \vec{a}_i^* are the same as after i follows \vec{a}_i' .*

PROOF. Fix any $h \in \mathcal{H}$ such that $\mu^{\vec{\sigma}^f}(h|h_i) > 0$. Fix two action profiles $\vec{a}^* = (\vec{a}_i^*, \vec{a}_{-i}^*)$ and $\vec{a}' = (\vec{a}_i', \vec{a}_{-i}')$, where \vec{a}_{-i}^* corresponds to an action profile prescribed by $\vec{\sigma}^*$ immediately after h , for every player other than i . Let $h^* = (h, \vec{a}^*)$ and $h' = (h, \vec{a}')$. By M4, i is either punished in every dissemination round after h^* and h' by every node or not punished by any node. If i is punished, then by D1, for any $z \in \mathcal{Z}^t$ succeeding h^* or h' , it holds $(id, e) \notin \text{rec}_i(z)$. To see this suppose the opposite. By construction, there exists j that sends $(id, g^h(e', i))$ to i in round r such that i is able to retrieve e from $g^h(e', i)$. However, since j is following σ_j^f in round r , it must hold $(id, e') \in PE_j$ after round $r - 1$. By D1, $e' = g^h(e, j)$. By Non-disclosure, i cannot retrieve e from $g^h(e', i)$, reaching a contradiction. Therefore, the expected benefits of receiving events introduced in the future is 0 when i is being punished after h .

Now, consider that i is not being punished after h . By D3, the probability of i receiving some tuple (id, e') when id is introduced after h is the same after h^* and h' . By D1, when some $j \neq i$ sends (id, e') to i , it holds $e' = g^h(g^h(e, j), i) = g^h(e, j)$, since $(id, g^h(e', i)) \in PE_j$ before j sends (id, e') to i . Consequently, i is able to retrieve e from e' , obtaining a benefit β with the same probability after h^* and h' . This implies that expected benefits are the same in this case. \square

Lemma 4.8 quantifies the difference of the expected costs of following two alternative actions \vec{a}_i^* and \vec{a}_i' , as a function of the states resulting from following these actions. Namely, we first consider the case where \vec{a}_i^* and \vec{a}_i' result in $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$ or $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$ and $\text{Miss}_i^* = \text{Miss}_i'$. This covers the scenarios where i sends only valid messages prescribed by the strategy, except the exact content in each message may vary. In addition, when $\Theta_i(i) = \text{Bad}$, this also covers scenarios where i sends invalid messages. Then, we analyse the case where, in \vec{a}_i^* , i forwards every event as specified by the strategy while, in \vec{a}_i' , i fails to forward some event ($\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$ and $\text{Miss}_i^* \subset \text{Miss}_i'$). This is representative of any scenario where \vec{a}_i^* is an action prescribed by the strategy, and \vec{a}_i' does not contain invalid

messages but also does not contain some tuple that i was supposed to forward. Finally, we analyse the case when i only sends valid messages in \vec{a}_i^* and sends some invalid message in \vec{a}_i' , when $\Theta_i(i)$ is initially *Good*.

LEMMA 4.8. *Fix any stage $t \in \mathbb{N}$, player $i \in \mathcal{N}$, history $h_i \in \mathcal{H}_i$ from stage t , and any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i$. Regarding the difference between the expected costs of forwarding events introduced after h_i is observed when i follows \vec{a}_i^* and when i follows \vec{a}_i' , we have:*

1. *If $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$ or $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$ and $\text{Miss}_i^* = \text{Miss}_i'$, then the difference is 0.*
2. *If $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$ and $\text{Miss}_i^* \subset \text{Miss}_i'$, then for every $\xi > 0$ there exists a function SG_i^f such that the difference is at least $-(q(f) + \xi)cn^E\gamma f/\nu$ for some constant $c > 0$.*
3. *If $\Theta_i^*(i) = \text{Good}$ and $\Theta_i'(i) = \text{Bad}$, then for every $\xi > 0$ there exists a function SG_i^f such that the difference is at least $-(q(f) + \xi)\gamma f$.*

PROOF. Fix any identifier $id \in \{1 \dots \nu\}$ introduced after h_i is observed.

Case 1. By Proposition 4.1 and the definition of $\vec{\sigma}^f$, i forwards id exactly once after its first reception, and only if $\Theta_i(i) = \text{Good}$. Thus, if $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$, then the expected costs are 0, whether i follows \vec{a}_i^* or \vec{a}_i' . If $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$, then by D3 i receives a tuple containing id after \vec{a}_i^* with the same probability as after \vec{a}_i' . By Proposition 4.1 and the fact that $\text{Miss}_i^* = \text{Miss}_i'$, i forwards id to exactly f nodes after \vec{a}_i^* iff i does so after \vec{a}_i' . Therefore, the expected costs are the same.

Case 2. First, suppose that $\text{seq}(id) \in \text{Miss}_i' \setminus \text{Miss}_i^*$. After following \vec{a}_i' , i does not forward id . After \vec{a}_i^* , by D7, i forwards id with probability at most $q(f) + \xi$ for an arbitrarily small ξ . Since there are at most n^E events per sequence and $|\text{Miss}_i'| - |\text{Miss}_i^*| \leq c$ for some constant c , this yields a difference of at least $-(q(f) + \xi)cn^E\gamma f/\nu$. Now, suppose $\text{seq}(id) \in \text{Miss}_i^* \cap \text{Miss}_i'$. By D3, i receives id after \vec{a}_i^* and \vec{a}_i' with the same probability, in which case i forwards id to exactly f nodes. Thus, the expected costs of forwarding id are the same.

Case 3. i does not forward id after \vec{a}_i' . By D7, i forwards id after \vec{a}_i^* with probability at most $q(f) + \xi$. Since there are at most ν events introduced after h_i , this difference is at least $-(q(f) + \xi)\nu\gamma f/\nu$. \square

Lemma 4.9 analyses the difference between the expected utility of player i following an action \vec{a}_i^* prescribed by the strategy for dissemination rounds, after the observation of h_i , and that of following an arbitrary action \vec{a}_i' . We consider three cases, which cover all the scenarios relevant to the proof that $\vec{\sigma}^f$ is an equilibrium. By Proposition 4.1, it holds $\Theta_i^*(i) = \Theta_i^{h_i}$ and $\text{Miss}_i^* = \text{Miss}_i^{h_i}$. Case 1 covers the scenarios where $\Theta_i^*(i) = \Theta_i^{h_i} = \text{Bad}$, where we always have $\Theta_i'(i) = \text{Bad}$, and covers the scenarios where $\Theta_i^*(i) = \text{Good}$ and i sends only valid messages in \vec{a}_i' and forwards every event as requested, implying that $\text{Miss}_i' = \text{Miss}_i^*$. Case 2 considers the alternative scenario where $\Theta_i^*(i) = \text{Good}$, i only forwards valid messages in \vec{a}_i' such that $\Theta_i'(i) = \text{Good}$, but i fails to send some identifier id such that $\text{seq}(id) \in \text{Miss}_i' \setminus \text{Miss}_i^*$. Finally, case 3 analyses the scenario where i sends some invalid message in \vec{a}_i' while $\Theta_i^*(i) = \text{Good}$.

LEMMA 4.9. Fix any player $i \in \mathcal{N}$, private history $h_i \in \mathcal{H}$ from stage $t \in \mathbb{N}$ preceding a dissemination round, and any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i$ such that $\sigma_i^f(\vec{a}_i^* | h_i) > 0$. The following holds regarding the value $\Delta^t = \mathbb{E}^{\vec{\sigma}^f, \mu, t}(u_i | h_i, \vec{a}_i^*) - \mathbb{E}^{\vec{\sigma}^f, \mu, t}(u_i | h_i, \vec{a}_i')$:

1. If $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$ or $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$ and $\text{Miss}_i^* = \text{Miss}_i'$, then $\Delta^t \geq 0$.
2. If $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$ and $\text{Miss}_i^* \subset \text{Miss}_i'$, then for every $\xi > 0$ there exists a function SG_i^f and a constant $c > 0$ such that:

$$\Delta^t \geq -\rho(\beta + \gamma f)/\nu - (q(f) + \xi)cn^E \gamma f/\nu.$$

3. If $\Theta_i^*(i) = \text{Good}$ and $\Theta_i'(i) = \text{Bad}$, then for every $\xi > 0$ there exists a function SG_i^f such that:

$$\Delta^t \geq -\rho(\beta + \gamma f)/\nu - (q(f) + \xi)\gamma f.$$

PROOF. Fix $h \in \mathcal{H}$ such that $\mu^{\vec{\sigma}}(h | h_i) > 0$, $\vec{a}^* = (\vec{a}_i^*, \vec{a}_{-i}^*)$, $\vec{a}' = (\vec{a}_i', \vec{a}_{-i}')$, $h^* = (h, \vec{a}^*)$, and $h' = (h, \vec{a}')$, such that \vec{a}_{-i}^* is prescribed by $\vec{\sigma}^f$ for every player other than i . We analyse each case separately.

Case 1. By Lemmas 4.7 and 4.8, expected costs and benefits relative to events introduced after h_i are the same. Now, we analyse costs and benefits of events introduced before h_i . Fix a tuple (id, e) introduced before h . If $\Theta_i^*(i) = \text{Bad}$, then i does not forward id in \vec{a}_i^* . Else, for any j , $id \in a_i^*(j)$ iff $\text{seq}(id) \notin \text{Miss}_i^h$, $j \in SG_i^f(sd_i, id)$, and i has received id for the first time in the previous round. In this case, it must hold $id \in a_i^*(j)$ only if $id \in a_i'(j)$. To see this, suppose that $id \in a_i^*(j)$ and $id \notin a_i'(j)$. This would imply $\text{seq}(id) \in \text{Miss}_i' \setminus \text{Miss}_i^*$, which would contradict case 1. By D5, i obtains a benefit β for receiving (id, e) after h' only if i obtains that benefit after h^* . Thus, expected benefits relative to (id, e) after i follows \vec{a}_i^* are at least as high as those after \vec{a}_i' .

The expected costs of forwarding id depend on whether i forwards id in \vec{a}_i^* and \vec{a}_i' , or forwards id after these actions are taken. We have seen that, for each $j \neq i$, $id \in a_i^*(j)$ only if $id \in a_i'(j)$. Thus, costs of following \vec{a}_i^* are at least as high as those of following \vec{a}_i' . First, suppose that $\text{seq}(id) \in \text{Miss}_i^* = \text{Miss}_i'$ or $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$. By Proposition 4.1, i does not forward id after h^* and h' . Therefore, expected costs of forwarding id after h^* and h' are both 0.

Now, consider that $\text{seq}(id) \notin \text{Miss}_i^* = \text{Miss}_i'$ and $\Theta_i'(i) = \text{Good}$. By Proposition 4.1 and the definition of $\vec{\sigma}^f$, i forwards id to f nodes after its first reception, both after h^* and h' . It must hold for every $j \neq i$ that $id \in a_i^*(j)$ iff $id \in a_i'(j)$. To see this, first suppose that $id \in a_i^*(j)$ and $id \notin a_i'(j)$ for some j . By definition of σ_i^f , $\text{seq}(id) \notin \text{Miss}_i^*$, which implies that i received id for the first time in the previous round and $j \in SG_i^f(sd_i, id)$. But in this case we would have $\text{seq}(id) \in \text{Miss}_i' \setminus \text{Miss}_i^*$. Now, suppose that $id \notin a_i^*(j)$ and $id \in a_i'(j)$. Again, $\text{seq}(id) \notin \text{Miss}_i^*$. Thus, either $j \notin SG_i^f(sd_i, id)$ or i did not receive id in the previous round. Either way, we would have $\text{seq}(id) \in \text{Miss}_i' \setminus \text{Miss}_i^*$. Both situations contradict the assumption that $\text{Miss}_i^* = \text{Miss}_i'$. Therefore, we can apply D4 to conclude that i receives id for the first time and forwards it after h^* with the same probability as after h' . Consequently, expected costs of forwarding id after h^* and h' are the same.

Cases 2 and 3. By Lemmas 4.7 and 4.8, the difference between the expected utilities for receiving and disseminating events introduced after h for cases 2 and 3 is at least $-(q(f) + \xi)cn^E \gamma f/\nu$ and $-(q(f) + \xi)\gamma f$, respectively. There are at most ρ events introduced in h . In the worst scenario, when following \vec{a}_i^* , i does not retrieve any of those events while forwarding all of them, resulting in the minimum expected difference of $-\rho(\beta + \gamma f)/\nu$. \square

In Lemma 4.9, we consider alternative actions \vec{a}_i^* and \vec{a}_i' followed in a monitoring round, where \vec{a}_i^* is prescribed by σ_i^f . Since we never update Miss_i in a monitoring round, we only need to consider the resulting values of $\Theta_i^*(i)$ and $\Theta_i'(i)$. By Proposition 4.1, we always have $\Theta_i^{h_i}(i) = \Theta_i^*(i)$, where h_i is the private history preceding \vec{a}_i^* . If $\Theta_i^*(i) = \text{Bad}$, then $\Theta_i'(i) = \text{Bad}$. Else, we have $\Theta_i'(i) = \text{Good}$ iff i sends all the required monitoring messages and only valid messages. These two scenarios are captured by case 1 of the lemma. Case 2 encompasses the other scenario where $\Theta_i^*(i) = \text{Good}$ and i sends some invalid message or fails to send a requested message in \vec{a}_i' .

LEMMA 4.10. Fix any player $i \in \mathcal{N}$, private history $h_i \in \mathcal{H}$ from stage $t \in \mathbb{N}$ preceding a monitoring round, and any two actions $\vec{a}_i^*, \vec{a}_i' \in \mathcal{A}_i$. The following holds regarding the value $\Delta^t = \mathbb{E}^{\vec{\sigma}^f, \mu, t}(u_i | h_i, \vec{a}_i^*) - \mathbb{E}^{\vec{\sigma}^f, \mu, t}(u_i | h_i, \vec{a}_i')$:

1. If $\Theta_i^*(i) = \Theta_i'(i)$, then $\Delta^t \geq 0$.
2. If $\Theta_i^*(i) = \text{Good}$ and $\Theta_i'(i) = \text{Bad}$, then for every $\xi > 0$ there exists a function SG_i^f such that:

$$\Delta^t \geq -n\alpha^*(1 + n^S p^*)/\nu - (q(f) + \xi)\gamma f.$$

PROOF. Fix $h \in \mathcal{H}$ such that $\mu^{\vec{\sigma}}(h | h_i) > 0$, $\vec{a}^* = (\vec{a}_i^*, \vec{a}_{-i}^*)$, $\vec{a}' = (\vec{a}_i', \vec{a}_{-i}')$, $h^* = (h, \vec{a}^*)$, and $h' = (h, \vec{a}')$, where \vec{a}_{-i}^* is prescribed by $\vec{\sigma}^f$ for every player other than i . We analyse each case separately.

Case 1. First, consider that $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$. By definition of σ_i^f and Proposition 4.1, i does not send messages in \vec{a}_i^* nor after h^* and h' . Thus, expected costs are at least as high when following \vec{a}_i^* as when following \vec{a}_i' . By Lemma 4.7, expected benefits are the same after h^* and h' . Now, suppose $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$. Since this is a monitoring round, it holds $\text{Miss}_i^* = \text{Miss}_i' = \emptyset$. By Lemmas 4.7 and 4.8, expected benefits and costs of receiving and forwarding events are the same after h^* and h' . By Proposition 4.1, i forwards every requested monitoring message after h^* and h' . Since mediator requests each report with fixed probability p^* , expected monitoring costs are the same. Finally, i sends at least as many messages in \vec{a}_i^* as in \vec{a}_i' , implying that $\Delta^t \geq 0$.

Case 2. Lemmas 4.7 and 4.8 imply that the contribution of the expected benefits and costs of dissemination to Δ^t is at least $-(q(f) + \xi)\gamma f$. Regarding monitoring, i sends in \vec{a}_i^* all requested messages, which cost at most $n\alpha^*$. By Proposition 4.1 and the definition of σ_i^f , i does not send any monitoring message after h' , and sends an expected number of $nn^S p^*$ messages after h^* : for each node j , i sends reports to the mediator per sequence of identifiers with independent probability p^* ; thus, i sends reports relative to j regarding an expected number of sequences $n^S p^*$. This adds at least the value $-n\alpha^*(1 + n^S p^*)/\nu$ to Δ^t . \square

4.3 Main Result

We now prove our main result. First, we establish sufficient conditions on the parameters n^E , n^S , and p^* . Then, we prove that $(\bar{\sigma}^f, \mu^f)$ is a SE. We conclude the section by proving our central claim.

4.3.1 Conditions on Parameters

The following conditions on n^E , n^S , and p^* are sufficient for our purposes:

C1. $n^E \log(\nu) = o(\nu)$.

C2. $p^* \log(\nu) = o(1)$.

C3. $p^* \nu = \Omega(n^E)$.

C4. For some constant $c > 0$, $\lim_{\nu \rightarrow \infty} (1 - p^*)^{n^S} = c$.

C1 and C2 are required for keeping the expected monitoring costs sub-linear on ν . C3 and C4 guarantee that the immediate gain of deviating does not out-weigh the future loss of being punished. One definition that fulfils all conditions is $n^E = n^S = \sqrt{\nu}$ and $p^* = 1/\sqrt{\nu}$.

4.3.2 Equilibria Existence

In Theorem 4.11, we show that, for any fanout $f > 0$, $(\bar{\sigma}^f, \mu^f)$ is Sequentially Rational, as long as δ and ν are sufficiently large and the benefit β per received event is sufficiently larger than the cost γf of forwarding each event to f nodes. Since $(\bar{\sigma}^f, \mu^f)$ is consistent by definition, it follows that $(\bar{\sigma}^f, \mu^f)$ is a SE. In the proof, we use the fact that $\alpha^* = 4\alpha n^E \log(\nu)$, shown as follows. α^* comprises the cost of i sending to the mediator $2n^E$ reports from $SE_i(j)$ and $RE_i(j)$. Each report has a round and event identifier, corresponding to $2 \log(\nu)$ bits.

THEOREM 4.11. *For every $f > 0$, there exist constants $c > 0$, $\bar{\delta} \in (0, 1)$, and $\bar{\nu} \in \mathbb{N}$, such that, for every $\delta \in (\bar{\delta}, 1)$ and $\nu > \bar{\nu}$, if $\beta > c\gamma f$, then $(\bar{\sigma}^f, \mu^f)$ is Sequentially Rational.*

PROOF. We use the One-deviation Property to show the result. More precisely, we fix any player $i \in \mathcal{N}$ and $h_i \in \mathcal{H}_i$ from any stage t , and we calculate, for any actions \bar{a}_i^* , $\bar{a}_i' \in \mathcal{A}_i$ such that $\sigma_i^f(\bar{a}_i^* | h_i) > 0$, the value:

$$\Delta = \mathbb{E}^{\bar{\sigma}^f, \mu^f}(u_i | h_i, \bar{a}_i^*) - \mathbb{E}^{\bar{\sigma}^f, \mu^f}(u_i | h_i, \bar{a}_i').$$

The goal is to prove that $\Delta \geq 0$ for ν and δ sufficiently large. We have $\Delta = \sum_{t'=t}^{\infty} \delta^{t'-t} \Delta^{t'}$, where:

$$\Delta^{t'} = \mathbb{E}^{\bar{\sigma}^f, \mu^f, t'}(u_i | h_i, \bar{a}_i^*) - \mathbb{E}^{\bar{\sigma}^f, \mu^f, t'}(u_i | h_i, \bar{a}_i').$$

By Lemma 4.5, it holds $\Delta^{t'} = 0$ for every $t' > t+1$, thus, we have $\Delta = \Delta^t + \delta \Delta^{t+1}$. We need to compute Δ for the cases where h_i precedes monitoring and dissemination rounds.

Monitoring Round. When $\Theta_i'(i) = \Theta_i^*(i)$, $\Delta \geq 0$ by 1 of Lemmas 4.6 and 4.10. Else, by 2 of the same lemmas, we have for an arbitrarily small $\xi > 0$:

$$\Delta \geq -n\alpha^*(1 + n^S p^*)/\nu - (q(f) + \xi)\gamma f + \delta(q(f) - \xi)\beta.$$

Since $n^S n^E = \nu$, by C1 and C2:

$$\begin{aligned} n\alpha^*(1 + n^S p^*)/\nu &= 4n\alpha n^E \log(\nu)(1 + n^S p^*)/\nu \\ &= O(p^* \log(\nu) + n^E \log(\nu)/\nu) = o(1). \end{aligned}$$

Therefore, for ν sufficiently large, there exists ϵ arbitrarily small such that:

$$\Delta \geq q(f)(\delta\beta - \gamma f) - \epsilon.$$

If $\beta > \gamma f$ and δ is sufficiently large, then $\Delta \geq 0$.

Dissemination Round. We need to consider the three cases where: 1) $\Theta_i^*(i) = \Theta_i'(i) = \text{Good}$ and $\text{Miss}_i^* = \text{Miss}_i'$ or $\Theta_i^*(i) = \Theta_i'(i) = \text{Bad}$; 2) $\Theta_i'(i) = \Theta_i^*(i) = \text{Good}$ and $\text{Miss}_i^* \subset \text{Miss}_i'$; and 3) $\Theta_i^*(i) = \text{Good}$ and $\Theta_i'(i) = \text{Bad}$.

Case 1. By 1) of Lemmas 4.6 and 4.9, $\Delta \geq 0$.

Case 2. By 2) of Lemmas 4.6 and 4.9, we have for some constant c and an arbitrarily small $\xi > 0$:

$$\begin{aligned} \Delta \geq & -o(\nu)/\nu - (q(f) + \xi)c(n^E/\nu)\gamma f + \\ & + \delta(q(f) - \xi)\beta(P(x') - P(x^*)), \end{aligned}$$

where $x^* = |\text{Miss}_i^*|$ and $x' = |\text{Miss}_i'|$. Since $x' - x^* \geq 1$:

$$\begin{aligned} & P(x') - P(x^*) \geq \\ & \geq (1 - (1 - p^*)^{x^*+1}) - (1 - (1 - p^*)^{x^*}) = \\ & = p^*(1 - p^*)^{x^*} \geq p^*(1 - p^*)^{n^S}. \end{aligned}$$

If ν is sufficiently large, then for an arbitrarily small ϵ :

$$\Delta \geq q(f)(\delta\beta p^*(1 - p^*)^{n^S} - c(n^E/\nu)\gamma f) - \epsilon.$$

If we divide both sides by $k = p^*(1 - p^*)^{n^S}$, then by C3 and C4, there exists $c' > 0$ such that for an arbitrarily small $\epsilon > 0$:

$$\Delta/k \geq q(f)(\delta\beta - c'\gamma f) - \epsilon.$$

If $\beta > c'\gamma f$, then for δ sufficiently close to 1 and ϵ sufficiently small we have $\Delta \geq 0$.

Case 3. By 3) of Lemmas 4.6 and 4.9, for an arbitrarily small $\xi > 0$:

$$\Delta \geq -o(\nu)/\nu - (q(f) + \xi)\gamma f + \delta(q(f) - \xi)\beta(1 - P(x^*)),$$

where $x^* = |\text{Miss}_i^*|$. Since $1 - P(x^*) \geq (1 - p^*)^{n^S}$, as in case 2), there exists a constant $c > 0$ such that for an arbitrarily small ϵ :

$$\Delta \geq q(f)(\delta\beta - c\gamma f) - \epsilon.$$

If $\beta > c\gamma f$, for δ sufficiently close to 1, it holds $\Delta \geq 0$. \square

4.3.3 Central Claim

We conclude with the proof of Theorem 2.3. The result follows trivially for $f = q(f) = 0$. Thus, fix any $f > 0$. We have already defined a monitoring mechanism and a pair $(\bar{\sigma}^f, \mu^f)$ that is a SE of the induced game, as long as δ and ν are sufficiently large, and $\beta = \Omega(\gamma f)$. By M1, M3, and D6, the expected utility of forwarding and receiving events per stage is $\bar{u}(f)$ fulfilling $|\bar{u}(f) - \bar{u}(f)| < \epsilon$ for an arbitrarily small ϵ . This yields an average utility of:

$$(1 - \delta) \sum_{t=0}^{\infty} \delta^t (-A/\nu + \bar{u}(f)) = -A/\nu + \bar{u}(f),$$

where $A \leq n\alpha^*(1 + p^* n^S)$. By C1 and C2, we have $A = o(\nu)$. For ν sufficiently large, A/ν is arbitrarily small, and the average utility is arbitrarily close to $\bar{u}(f)$.

5. DISCUSSION

The main step towards improving our result lies in distributing the role of the mediator, possibly having a different mediator per node. This goal poses two main challenges: 1) how to ensure that every node agrees on whether to punish any i and on the key κ_i ; and 2) how to distribute the

seeds. The first challenge can be addressed by having nodes reaching an agreement regarding the set of punishments and keys for each stage, in a similar fashion to [8]. The second challenge can be addressed using techniques similar to those from [2], by having nodes committing to random numbers without revealing each seed sd_i to any node other than i .

Acknowledgements

This work was partially supported by Fundação para a Ciência e Tecnologia (FCT) via the INESC-ID multi-annual funding through the PIDDAC Program fund grant, under project PEst-OE/EEI/LA0021/2013, and via the project PEPITA (PTDC/EEI-SCR/2776/2012).

6. REFERENCES

- [1] H. Abbey. An examination of the Reed-Frost theory of epidemics. *Human Biology*, 24(3):201–233, 1952.
- [2] I. Abraham, D. Dolev, and J. Halpern. Distributed protocols for leader election: A game-theoretic perspective. In *DISC'13*, pages 61–75. Springer-Verlag, 2013.
- [3] S. Aiyer, L. Alvisi, A. Clement, M. Dahlin, J. Martin, and C. Porth. BAR fault tolerance for cooperative services. In *SOSP'05*, pages 45–58. ACM, 2005.
- [4] R. Axelrod. *The evolution of cooperation*. Basic Books, 1984.
- [5] M. Bellare, A. Desai, E. Jorjani, and P. Rogaway. A concrete security treatment of symmetric encryption. In *FOCS'97*, pages 394–. IEEE Computer Society, 1997.
- [6] K. P. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, and Y. Minsky. Bimodal multicast. *ACM Trans. Comput. Syst.*, 17(2):41–88, 1999.
- [7] B. Cohen. Incentives build robustness in bittorrent. In *P2PEcon'03*, 2003.
- [8] S. Dolev, E. Schiller, P. Spirakis, and P. Tsigas. Game authority for robust and scalable distributed selfish-computer systems. *Theor. Comput. Sci.*, 411(26–28):2459–2466, 2010.
- [9] R. Guerraoui, K. Huguenin, A. Kermarrec, M. Monod, and S. Prusty. Lifting: lightweight freerider-tracking in gossip. In *Middleware'10*, pages 313–333, 2010.
- [10] E. Hendon, H. Jacobsen, and B. Sloth. The one-shot-deviation principle for sequential rationality. *Games and Economic Behavior*, 12(2):274–282, 1996.
- [11] ISO/IEC. Information technology – Security techniques – Modes of operation for an n-bit block cipher. ISO ISO/IEC 10116:2006, International Organization for Standardization, Geneva, Switzerland, 2006.
- [12] I. Keidar, R. Melamed, and A. Orda. Equicast: Scalable multicast with selfish users. *Comput. Netw.*, 53(13):2373–2386, 2009.
- [13] D. Kreps and R. Wilson. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982.
- [14] H. C. Li, A. Clement, M. Marchetti, M. Kapritsos, L. Robison, L. Alvisi, and M. Dahlin. Flightpath: obedience vs. choice in cooperative services. In *OSDI'08*, pages 355–368, 2008.
- [15] H. C. Li, A. Clement, E. L. Wong, J. Napper, I. Roy, L. Alvisi, and M. Dahlin. BAR gossip. In *OSDI'06*,

pages 191–204, 2006.

- [16] G. Mailath and L. Samuelson. Repeated games and reputations. Oxford University Press, Oxford, 2007.
- [17] M. Osborne and A. Rubinstein. *A course in game theory*. The MIT Press, 1994.
- [18] R. Trivers. The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1):35, 1971.
- [19] X. Vilaca and L. Rodrigues. On the effectiveness of punishments in a repeated epidemic dissemination game. In *SSS'13*, pages 206–220. Springer-Verlag, 2013.
- [20] E. Wong, J. Leners, and L. Alvisi. It's on me! the benefit of altruism in BAR environment. In *DISC'10*, pages 406–420. Springer, 2010.

APPENDIX

A. CORRECTNESS PROOFS

Proof of Lemma 4.3

D1. We use induction to show that the property holds for every history $h' = h$ or succeeding h , and preceding h^* . For the base case, we have $h' = h$. Since id has not been disseminated prior to h , no node j has id in PE_j . For the induction step, suppose that D1 holds for some history h' , and fix any j . If $(id, e') \in PE_j^{h'}$, then the result follows by the hypothesis. If j is the source and disseminates id when h' is observed, then only j adds (id, e) to PE_j after h' . Else, consider that j receives (id, e') from $l \in \mathcal{N}$. By the hypothesis, we have $e' = g^h(g^h(e, l), j)$, and j adds (id, e'') to PE_j with $e'' = g^h(e', l)$. By Commutativity, $e'' = g^h(e, j)$.

D2. Let $\vec{a}^* = (\vec{a}_i^*, \vec{a}_{-i})$ and $\vec{a}' = (\vec{a}'_i, \vec{a}_{-i})$ be the action profiles followed after h when players different from i abide to $\vec{\sigma}^f$. Fix any two histories h^* and h' from stage t' , immediately preceding the last monitoring round of stage t' , and succeeding (h, \vec{a}^*) and (h, \vec{a}') , respectively. By the fact that we initialise $\Theta_j(j) = Good$ and $Miss_j = \emptyset$ in stage t' , and by Proposition 4.1, it holds $\Theta_j^{h^*}(j) = \Theta_j^{h'}(j) = Good$ and $Miss_j^{h^*} = Miss_j^{h'} = \emptyset$ for every $j \in \mathcal{N}$. Thus, j abides to $\vec{\sigma}^f$ throughout stage t' by forwarding a tuple containing id once after its first reception, to the f nodes from $SG_j^f(sd_j, id)$. Thus, if the selection of seeds after h^* is the same as after h' , then i receives a tuple containing id after h^* iff i receives id after h' . Since the probability distribution of the selection of seeds is the same after h^* and h' , the result follows.

D3. Let $\vec{a}^* = (\vec{a}_i^*, \vec{a}_{-i})$ and $\vec{a}' = (\vec{a}'_i, \vec{a}_{-i})$ be the action profiles followed after h when players different from i abide to $\vec{\sigma}^f$. First, suppose that h precedes a dissemination round. Since the mediator is trusted, regardless of the action followed by i , the value of sd_j is fixed for every $j \in \mathcal{N}$. If $j \neq i$, then by Proposition 4.1 whether i follows \vec{a}_i^* or \vec{a}'_i the values $\Theta_j(j)$ and $Miss_j$ in any dissemination round are equal to $\Theta_j^h(j)$ and $Miss_j^h$, respectively. Therefore, the behaviour of j after i follows \vec{a}_i^* is the same as after i follows \vec{a}'_i : j either forwards id after its reception according to $SG_j^f(sd_j, id)$, or j does not forward id to any node. Thus, i receives a tuple containing id after following \vec{a}_i^* iff i receives id after following \vec{a}'_i . Now, suppose that h precedes a monitoring round and let h^* and h' be any histories immediately preceding the last monitoring round and succeeding (h, \vec{a}^*) and (h, \vec{a}') , respectively. As in Property D2, for the same

selections of seeds after h^* and h' , i receives id after h^* iff i receives id after h' . Since the probability distribution over the seeds is the same, i receives id with the same probability.

D4. Let $\vec{a}^* = (\vec{a}_i^*, \vec{a}_{-i})$ and $\vec{a}' = (\vec{a}_i', \vec{a}_{-i}')$ be the action profiles followed after h when players different from i abide to $\vec{\sigma}^f$. i receives id in (h, \vec{a}^*) iff the same holds in (h', \vec{a}') . Now, we show using induction that, for every $j \neq i$ and histories h^* and h' succeeding or equal to (h, \vec{a}^*) and (h, \vec{a}') respectively: 1) j receives id in h^* iff j receives id in h' ; and 2) j forwards id immediately after h^* iff j forwards id immediately after h' . This proves that i receives id after \vec{a}^* iff i receives id after \vec{a}' .

Base case. 1) holds if j receives id in h . Otherwise, j receives id in \vec{a}^* iff j receives id in \vec{a}' , proving 1). By construction, $\Theta_j^{h^*}(j) = \Theta_j^{h'}$ and $Miss_j^{h^*} = Miss_j^{h'}$. This means that j has $\Theta_j^{h^*}(j) = Bad$ or $seq(id) \in Miss_i^{h^*}$ iff $\Theta_j^{h'}(j) = Bad$ or $seq(id) \in Miss_i^{h'}$. This implies that j forwards id immediately after h^* iff the same holds after h' , proving 2).

Induction step. Consider that the hypothesis holds for any h^* and h' . It holds that j receives id in h^* iff j receives id in h' , implying 1). In this case, j does not forward id one round after h^* and h' , which implies 2). Now, by 2) of the hypothesis, j first receives id from some l after h^* iff j first receives id from l after h' , proving 1). By Proposition 4.1, $\Theta_j^{h^*}(j) = \Theta_j^{h'}(j)$ and $Miss_j^{h^*} = Miss_j^{h'}$, proving 2).

D5. Fix any z^* and z' . Let $\vec{a}^* = (\vec{a}_i^*, \vec{a}_{-i})$ and $\vec{a}' = (\vec{a}_i', \vec{a}_{-i}')$ be the action profiles immediately following h in z^* and z' , respectively. i receives (id, e') in (h, \vec{a}^*) iff i receives (id, e') in (h, \vec{a}') , whether i can retrieve e from e' . Thus, if $(id, e) \in rec_i(z')$ because i received (id, e') in (h, \vec{a}') , then it also holds $(id, e) \in rec_i(z^*)$.

Now, suppose that for all (id, e') received in (h, \vec{a}') and (h, \vec{a}^*) it holds that i cannot retrieve e from e' . Suppose also that $(id, e) \in rec_i(z')$. Here, some $j \neq i$ sends (id, e') to i after (h, \vec{a}') such that i can retrieve e from e' or $g^h(e', j)$. This occurs after some history \bar{h} , such that $(id, e') \in PE_j^{\bar{h}}$ with $e' = g^h(e'', i)$, and i is able to retrieve e from $g^h(g^h(e', j), i)$. In this case, we say that i can retrieve e from e' in j .

We prove two sub-properties. First, we show using induction that, for any histories h' and h^* succeeding or equal to (h, \vec{a}') and (h, \vec{a}^*) respectively, any $j \neq i$ forwards id immediately after h^* only if j forwards id immediately after or in h' . Then, we use this fact to prove through induction that, for any histories h' and h^* succeeding (h, \vec{a}') and (h, \vec{a}^*) respectively, and any $j \neq i$, if i can retrieve e from e' in j then: 1) $(id, e') \in PE_j^{h'}$ only if $(id, e') \in PE_j^{h^*}$; and 2) j forwards (id, e') one round after h' only if j forwards (id, e') one round after h^* . Given that these properties hold, i receives (id, e') from any j after (h, \vec{a}') only if i receives (id, e') from j after (h, \vec{a}^*) . Thus, if $(id, e) \in rec - i(z')$, then $(id, e) \in rec_i(z^*)$.

Property 1. Base case. It is true that if j has received id in h , then j never forwards id after (h, \vec{a}^*) and (h, \vec{a}') . In addition, j receives id in \vec{a}^* for the first time only if j receives id in \vec{a}' for the first time. That is, j receives id from any $l \neq i$ in \vec{a}^* iff the same holds in \vec{a}' , whereas j never receives id in \vec{a}^* from i while not receiving it from i in \vec{a}' . By the preservation of $\Theta_j(j)$ and $Miss_j$, it holds that j forwards id immediately after \vec{a}^* only if j forwards id after \vec{a}' .

Induction step. Suppose the hypothesis holds for two histories h^* and h' . Let $h^1 = (h^*, \vec{a}^1) \in z^*$ and $(h', \vec{a}^2) \in z'$ be the histories immediately succeeding h^* and h' , respec-

tively. Fix any $j \neq i$. If j already received and forwarded id in h^1 , then by the hypothesis j already forwarded id in h^2 , and does not forward id after h^2 . Now, consider that j receives id for the first time immediately after h^* , and forwards id immediately after h^1 . By construction, it must hold that $\Theta_j^{h^1}(j) = Good$ and $seq(id) \notin Miss_j^{h^1}$. By Proposition 4.1, $\Theta_j^{h^1}(j) = \Theta_j^{h^2}(j)$ and $Miss_j^{h^1} = Miss_j^{h^2}$. By the hypothesis, either j forwarded id to j in h^2 or j receives id immediately after h' . In the former case, the result follows. In the latter case, since we also have $\Theta_j^{h^2}(j) = Good$ and $seq(id) \notin Miss_j^{h^2}$, j forwards id immediately after h^2 . This proves the result.

Property 2. Base case. Suppose j receives (id, e'') from l with $e'' = g^h(e', l)$, such that l is the node with the smallest identifier among the nodes sending id to i in \vec{a}' , and adds (id, e') to $PE_j^{h'}$. We cannot have $l = i$, or else i sends (id, e'') to j before receiving this tuple and such that i can retrieve e from e'' , implying that $(id, e) \notin rec_i(z')$. Therefore, l also sends (id, e'') to j in \vec{a}^* . The set of nodes sending (id, e'') to j in \vec{a}^* is a subset of those that send id in \vec{a}' . Thus, j also receives (id, e'') from l , l also has the smallest identifier, and j adds (id, e') to $PE_j^{h^*}$, proving the base case.

Induction step. Suppose the hypothesis holds for h' and h^* . Let $h^1 = (h^*, \vec{a}^1) \in z^*$ and $(h', \vec{a}^2) \in z'$ be the histories immediately succeeding h^* and h' , respectively. Fix any $j \neq i$ and e' such that i can retrieve e from e' in j . By the hypothesis, if $(id, e') \in PE_j^{h'}$, then $(id, e') \in PE_j^{h^*}$; if j forwards id immediately after h' , then j forwards id immediately after h^* . Either way, j does not forward id after h^1 and h^2 , proving the result. Now, suppose that $(id, e') \notin PE_j^{h'}$ and that j receives id for the first time immediately after h' from a set of nodes S' . Let $l \in S'$ be the node with the smallest identifier, which may be i .

If l sends (id, e') with $e'' \neq g^h(e', l)$, then j never adds (id, e') to PE_j and never forwards this tuple after h' . Otherwise, $(id, e') \in PE_j^{h^2}$ and $(id, e^1) \in PE_l^{h'}$, where $e^1 = g^h(e'', j)$. Since i can retrieve e from e' in j , i can retrieve e from e^1 in l : i can retrieve e from $g^h(e^2, i)$ where $e^2 = g^h(e', j)$; by Commutativity, we have $e^1 = g^h(g^h(e', l), j) = g^h(e^2, l)$; thus, i can retrieve e from $g^h(g^h(e^1, l), i) = g^h(e^2, i)$. Therefore, we can apply 2 of the hypothesis to conclude that $(id, e^2) \in PE_l^{h^*}$, and l also sends (id, e'') immediately after h^* to j . Let S^* be the set of nodes sending id to j immediately after h^* . It must hold that j receives id for the first time after h^* , or else we would reach a contradiction with Property 1): if j had received id in h^* , then he would also have received id in h' , contradicting the supposition that j receives id for the first time after h' . It thus follows from Property 1) that, if any node x sends id to i immediately after h^* , then x sends id immediately after h' , i.e., $S^* \subseteq S'$. Since $l \in S^*$, l is the node with the smallest identifier in S^* . Thus, j adds (id, e') to $PE_j^{h^1}$, proving 1). In this case, by Proposition 4.1, $Miss_j^{h^*} = Miss_j^{h'}$ and $\Theta_j^{h^*}(j) = \Theta_j^{h'}(j)$, and j forwards id immediately after h^1 iff j forwards id after h^2 . This proves the result.

Proof of Lemma 4.4 Properties D6) and D7) follow from Lemmas C.1 and C.2, respectively, proven in Appendix C.

B. PSEUDO RANDOM NUMBER GENERATOR

We consider a pseudo-random number generator function defined as follows. The definition used in [2] considers a pseudo-random number generator function G defined as a sequence of $m(k) > k$ bits generated from a seed $s \in \{0, 1\}^k$ for some security parameter k . G is said to be a pseudo-random number generator if no probabilistic time machine can distinguish between the outcome of G and truly random sequences of bits. We formalise this intuition in a way that is useful for our purposes as follows. For all $\xi > 0$ and polynomial m , there exists k such that for every sequence of bits $\vec{b} \in \{0, 1\}^{m(k)}$:

$$\left| \frac{1}{2^{m(k)}} - \frac{|S|}{2^k} \right| < \xi, \quad (6)$$

where S is the largest set of seeds such that for every $s \in S$ we have $\vec{b} = G(s)$.

For each $i \in \mathcal{N}$, we define SG_i^f from G as follows. Let $x = \binom{n-1}{f}$ be the number of different subsets of nodes to whom i may send any event. We let $m(k) = \nu \cdot \phi(k) \cdot \lceil \log(x) \rceil$ for some polynomial $\phi(k)$. Any generated stream represents a sequence of ν numbers. Given the id -th number b , $SG_i^f(sd_i, id)$ returns the y -th subset, where $y \equiv b \pmod{x}$. By the definition of G , the probability of selecting a particular sequence of numbers is arbitrarily close to $1/x^\nu$, shown as follows. By using the modulo operation, we ensure that there exist two integers (d, r) such that: i) $r \in \{0 \dots x-1\}$; ii) $m(k) = xd + r$; iii) r subsets of each id are mapped into $d+1$ numbers; and iv) $x-r$ subsets are mapped into d numbers. The probability of each sequence of subsets being selected is arbitrarily close to selecting for each id a number that yields the corresponding subset. Since for each id and subset, there are d or $d+1$ different numbers yielding that subset, by 6, the probability of selecting each subset is either arbitrarily close to $d/m(k)$ or $(d+1)/m(k)$. Thus, the probability of selecting the sequence is arbitrarily close to a value lying in the interval:

$$\left(\left(\frac{d}{xd+r} \right)^\nu, \left(\frac{d+1}{xd+r} \right)^\nu \right).$$

By taking the limit $\phi(k) \rightarrow \infty$, we also have $d \rightarrow \infty$, and the probability of each sequence of subsets being selected converges to $1/x^\nu$. Thus, SG_i^f fulfils Property PRNG1 as shown by Proposition B.1. Here, we can adjust the constant ξ by adjusting k and $\phi(k)$.

PROPOSITION B.1. *Fix any node $i \in \mathcal{N}$ and sequence of subsets $\vec{S} = (S_{id})_{id \in \{1 \dots \nu\}}$ with $S_{id} \subseteq \mathcal{N} \setminus \{i\}$ and $|S_{id}| = f$ for every $id \in \{1 \dots \nu\}$. For any constant $\xi > 0$, there exist k and $\phi(k)$ such that for every $id \in \{1 \dots \nu\}$:*

$$\left| P^{SG_i^f}(\vec{S}|\vec{S}_{-id}) - \frac{1}{\binom{n-1}{f}} \right| \leq \xi.$$

PROOF. Fix any $\xi > 0$. By the definition of SG_i^f , we can define k and $\phi(k)$ such that:

$$|1/x^\nu - P^{SG_i^f}(\vec{S})| < \xi, \quad (7)$$

where $P^{SG_i^f}(\vec{S}) = \#S/2^k$ is the probability of selecting a seed such that SG_i^f yields \vec{S} . We can write:

$$P^{SG_i^f}(\vec{S}|\vec{S}_{-id}) = a/b = (a/2^k)/(b/2^k) = P^{SG_i^f}(\vec{S})/P^{SG_i^f}(\vec{S}_{-id}),$$

where a is the number of seeds that yield \vec{S} , and b is the number of seeds that yield \vec{S}_{-id} , independently of S_{id} . In particular, by 6, we can define SG_i^f such that:

$$\begin{aligned} P^{SG_i^f}(\vec{S}_{-id}) &= \\ &= \sum_{S'_{id} \subseteq \mathcal{N} \setminus \{i\}: |S'_{id}|=f} P^{SG_i^f}(S'_{id}, \vec{S}_{-id}) > \\ &> \sum_{S'_{id} \subseteq \mathcal{N} \setminus \{i\}: |S'_{id}|=f} (1/x^\nu - \xi) = x * (1/x^\nu - \xi). \end{aligned}$$

Similarly:

$$P^{SG_i^f}(\vec{S}_{-id}) < x * (1/x^\nu + \xi).$$

Therefore, we can write:

$$\begin{aligned} P^{SG_i^f}(\vec{S})/P^{SG_i^f}(\vec{S}_{-id}) &> \\ &> (1/x^\nu - \xi)/(x(1/x^\nu + \xi)) = \\ &= 1/(x(1 + x^\nu \xi)) - \xi x^\nu/(x(1 + x^\nu \xi)) = \\ &= 1/x - 2\xi x^\nu/(x(1 + x^\nu \xi)). \end{aligned}$$

Thus, we can define SG_i^f such that for an arbitrarily small $\xi' > 0$ it holds $P^{SG_i^f}(\vec{S})/P^{SG_i^f}(\vec{S}_{-id}) > 1/x - \xi'$. Using an identical reasoning, we can also define k and $\phi(k)$ such that $P^{SG_i^f}(\vec{S})/P^{SG_i^f}(\vec{S}_{-id}) < 1/x + \xi'$. This shows that $|P^{SG_i^f}(\vec{S}|\vec{S}_{-id}) - 1/x| < \xi'$ for an arbitrarily small ξ' , as we intended to prove. \square

C. DISSEMINATION MODEL

For any fanout f , the reliability of dissemination $q(f)$ is the probability of any node i receiving an event, and is determined using a model similar to the SIR model [1]. More precisely, the source introduces $e \in \mathcal{E}^t$ with identifier id in round $r^* = id + \tau^M - 1$ of stage t . In every round r , $Q^{id,f}(S, I, R, r)$ is the probability that nodes from the set S have received and forwarded a message containing id by round r , I contains the nodes that received a message containing id and will forward it in the next round, and R contains the nodes that have not received id , such that (S, I, R) forms a partition of \mathcal{N} . This probability is defined recursively as $Q^{id,f}(\emptyset, \emptyset, \mathcal{N}, r) = 1$ for $r < r^*$, $Q^{id,f}(\emptyset, \{s\}, \mathcal{N} \setminus \{s\}, r^*) = 1$, where s is the source, and for any $r > r^*$, partition (S, I, R) , and $I' \subseteq R$:

$$\begin{aligned} &Q^{id,f}(S, I, R, r+1) = \\ &= \sum_{I' \subseteq S} Q^{id,f}(S \setminus I', I', R \cup I, r) P^{id,f}(I|I', R), \end{aligned}$$

where $P^{id,f}(I|I', R)$ is the probability of exactly all nodes in I among $R \cup I'$ receiving a dissemination message containing id from some node in I' , given that each node forwards id to a subset of f neighbours chosen uniformly at random. The dissemination ends in round $r = \tau^M + id + \rho$. Then, $q(f)$ is the probability of any node i receiving id during the present stage and is defined as:

$$q(f) = \sum_{S \subseteq \mathcal{N}: i \in S} Q^{id,f}(S, \emptyset, \mathcal{N} \setminus S, r). \quad (8)$$

C.1 Auxiliary Proofs

We show that disseminating events using the pseudo-random number generator approximates the dissemination model for any fanout f . Namely, for each partition (S, I, R) of \mathcal{N} , we denote by $\mathcal{H}^{id,t}(S, I, R, r)$ the set of histories h of stage t such that $rn^h = r$ and S equals to the set of nodes that already received id , while I is the set of nodes that will forward id

in the next round:

$$S = \{j \in \mathcal{N} \mid \exists_{r'} \exists_{o \in \mathcal{N} \setminus \{j\}} (id, r') \in RE_j^h(o)\}.$$

$$I = \{j \in \mathcal{N} \setminus S \mid \exists_{e \in \mathcal{E}} (id, e) \in PE_j^h\}.$$

Lemma C.1 demonstrates that every player i expects every event to be disseminated in an approximate fashion to the dissemination model corresponding to any fanout, in future stages.

LEMMA C.1. *Fix any fanout f , stage $t \in \mathbb{N}$, terminal history $z \in \mathcal{Z}^t$, player $i \in \mathcal{N}$, and identifier $id \in \{1 \dots \nu\}$. For every $r \in \{\tau^M + id - 1 \dots \tau\}$, $\xi > 0$, and partition (S, I, R) of \mathcal{N} such that $i \in R$, there exists a function SG_i^f such that:*

$$\left| \sum_{h \in \mathcal{H}^{id, t+1}(S, I, R, r)} P^{\bar{\sigma}^f}(h|z) - Q^{id, f}(S, I, R, r) \right| \leq \xi. \quad (9)$$

PROOF. We show the hypothesis using induction on r .

Base case. We have $r = \tau^M + id - 1$. Fix any $h \in \mathcal{H}$ such that $P^{\bar{\sigma}^f}(h|z) > 0$ with $rn^h = r$. By construction, we have $(id, e) \in PE_s^h$, with s being the source and e the event disseminated with identifier id . Since no other node disseminates events, it holds that $\mathcal{H}^{id}(\emptyset, \{s\}, \mathcal{N} \setminus \{s\}, r)$ includes every history leading to round r . Thus, it holds:

$$\sum_{h \in \mathcal{H}^{id, t}(\emptyset, \{s\}, \mathcal{N} \setminus \{s\}, r)} P^{\bar{\sigma}^f}(h|z) = 1 = Q^{id, f}(\emptyset, \{s\}, \mathcal{N} \setminus \{s\}, r).$$

This proves the base case.

Induction step. Suppose that the hypothesis holds for some r . By construction, we have $\Theta_i^h(i) = \text{Good}$ and $\text{Miss}_i^h = \emptyset$ for every $i \in \mathcal{N}$ and history h with $rn^h = r$. Fix any history h^1 immediately preceding the last monitoring round of stage $t+1$. For any partition (S, I, R) and $h \in \mathcal{H}^{id, t}(S, I, R, r)$ succeeding h^1 , there exists exactly one history h^* per set $I' \subseteq R$ succeeding h where exactly the nodes from I' receive id from some node in I after h :

$$h^* \in \mathcal{H}^{id, t}(S \cup I, I', R \setminus I', r+1).$$

For each such I' , it holds:

$$\cup_{j \in I} SG_j^f(sd_j^h, id) \setminus S = I'. \quad (10)$$

Since seeds are determined in the last monitoring round, if we add over every $h \in \mathcal{H}^{id}(S, I, R, r)$ where nodes of I to exactly select nodes from I' , then we are adding over all the seeds selected by the mediator for nodes in I after h^1 , such that 10 holds. By PRNG1, when every seed is generated independently and uniformly at random, the probability of such even occurring is arbitrarily close to that when nodes in I select the subset of neighbours to whom they forward the event uniformly at random. The latter event occurs with probability $P^{id, f}(I'|I, R)$. That is, an approximate fraction $P^{id, f}(I'|I, R)$ of the histories from $\mathcal{H}^{id, t}(S, I, R, r)$ has nodes from I selecting those from I . By the hypothesis, for each

I' , there exists SG_i such that:

$$\begin{aligned} & \sum_{h \in \mathcal{H}^{id, t}(S \cup I, I', R \setminus I', r+1)} P^{\bar{\sigma}^f}(h|h^1) \approx \\ & \approx \sum_{h \in \mathcal{H}^{id, t+1}(S, I, R, r)} P^{\bar{\sigma}^f}(h|h^1) P^{id, f}(I'|I, R) \approx \\ & \approx Q^{id, f}(S, I, R, r) P^{id, f}(I'|I, R) = \\ & = Q^{id, f}(S \cup I, I', R \setminus I', r+1). \end{aligned}$$

Since this holds for any h^1 , the result follows. \square

LEMMA C.2. *Fix any fanout f , stage $t \in \mathbb{N}$, player $i \in \mathcal{N}$, private history $h_i \in \mathcal{H}_i$, and identifier $id \in \{1 \dots \nu\}$ not yet introduced by the source in h_i . For every $r \in \{\tau^M + id - 1 \dots \tau\}$, $\xi > 0$, and partition (S, I, R) of \mathcal{N} such that $i \in R$, there exists a function SG_i^f such that:*

$$\sum_{h \in \mathcal{H}^{id, t}(S, I, R, r)} P^{\bar{\sigma}^f, \mu}(h|h_i) - Q^{id, f}(S, I, R, r) \leq \xi. \quad (11)$$

PROOF. The proof is identical to Lemma C.1, except that now i has acquired some knowledge regarding the outputs of SG_j^f for each $j \neq i$, after the observation of h_i .

Base case. Identical to Lemma C.1.

Induction step. Suppose that the hypothesis holds for some r . For any partition (S, I, R) , fix any $h \in \mathcal{H}^{id, t}(S, I, R, r)$ such that $P^{\bar{\sigma}^f, \mu}(h|h_i) > 0$, and fix any $I' \subseteq R$. Fix $\vec{S} = (S_{id'}^j)$, with one subset $S_{id'}^j$ per node $j \neq i$ and identifier id' such that $|S_{id'}^j| = f$ and $S_{id'}^j = SG_j^f(sd_j^h, id')$. Fix h^1 preceding h and immediately preceding the last monitoring round from stage t . Since the actions observed between h^1 and h do not depend on $SG_j^f(sd_j^h, id)$, for each $(L_{id}^j)_{j \in I}$ with $|L_{id}^j| = f$, there exists a history $h' \in \mathcal{H}^{id, t}(S, I, R, r)$ such that $P^{\bar{\sigma}^f, \mu}(h|h^1) = P^{\bar{\sigma}^f, \mu}(h'|h^1)$ and h' may differ from h exactly in that $SG_j^f(sd_j^{h'}, id) = L_{id}^j$ for each $j \in I$. In particular, this holds for any history h' such that every node in I' receives id from some node in I , i.e.:

$$\cup_{j \in I^*} SG_j^f(sd_j^{h'}, id) \setminus S = I', \quad (12)$$

where $I^* \subseteq I$ is the largest subset such that for every $j \in I^*$ we have $\Theta_j^h(j) = \text{Good}$ and $\text{seq}(id) \notin \text{Miss}_j^h$. If we add over every such history h' , then we are adding over all the seeds for nodes in I^* such that 12 holds. By PRNG1, when every seed among the set that yields $S_{id'}^j$ for every $id' \neq id$ is generated uniformly at random, this is approximately close to nodes in I^* selecting the subset of neighbours to whom they forward the event uniformly at random, with an error lower than ξ . This event occurs with probability $P^{id, f}(I'|I^*, R) \leq P^{id, f}(I'|I, R)$, since $I^* \subseteq I$. For every such h' , the history immediately succeeding belongs to $\mathcal{H}^{id, t}(S \cup I, I', R \setminus I', r+1)$. Since the above holds for any h , an approximate fraction $P^{id, f}(I'|I^*, R)$ of the histories from $\mathcal{H}^{id, t}(S, I, R, r)$ has the nodes from I^* forwarding id to those from I' . Hence, we can write:

$$\begin{aligned} & \sum_{h \in \mathcal{H}^{id, t}(S \cup I, I', R \setminus I', r+1)} P^{\bar{\sigma}^*}(h|z) \leq \\ & \leq \sum_{h \in \mathcal{H}^{id, t}(S, I, R, r)} P^{\bar{\sigma}^*}(h|z) (P^{id, f}(I'|I^*, R) + \xi) \leq \\ & \leq \sum_{h \in \mathcal{H}^{id, t}(S, I, R, r)} P^{\bar{\sigma}^*}(h|z) (P^{id, f}(I'|I, R) + \xi). \end{aligned}$$

By the hypothesis, for each I' , there exists SG_i^f such that for an arbitrarily small $\xi' > 0$:

$$\begin{aligned}
& \sum_{h \in \mathcal{H}^{id,t}(S \cup I, I', R \setminus I', r+1)} P^{\vec{\sigma}^*}(h|z) && \leq \\
& \leq (Q^{id,f}(S, I, R, r) + \xi')(P^{id,f}(I'|I, R) + \xi) && = \\
& \leq Q^{id,f}(S, I, R, r)P^{id,f}(I'|I, R) + && \\
& \quad + \xi'(\xi + P^{id,f}(I'|I, R) + Q^{id,f}(S, I, R, r)) && = \\
& = Q^{id,f}(S \cup I, I', R \setminus I', r+1) + \xi'', &&
\end{aligned}$$

where:

$$\xi'' = \xi'(\xi + P^{id,f}(I'|I, R) + Q^{id,f}(S, I, R, r)).$$

Since ξ'' is also arbitrarily small, the result holds. \square